

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 11:13:02 ; Search time 17522 Seconds
(without alignments)
11673.789 Million cell updates/sec

Title: US-10-023-649-1_COPY_1_5000
Perfect score: 5000
Sequence: 1 accctacaactacgacgc.....ataatctagaacgcacc 5000

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_srs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pac:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_srs:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_ey:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5000	100.0	11484	14 AF214040	AF214040 Western e
2	3536	70.7	4471	14 WEU01065	U01065 Western equ
3	3536	70.7	4498	14 WEVNS	X74892 Western equ
4	2766	55.3	11678	14 U01034	U01034 Eastern equ
5	2745.2	54.9	11675	14 EEEVIRNA	X61335 Eastern Equ
6	2133.4	42.7	11441	14 VEUS5362	US5362 Venezuelan
7	2128.6	42.6	11420	14 AF004472	AF004472 Venezuela
8	2123.8	42.5	11420	14 AF004458	AF004458 Venezuela
9	2122.2	42.4	11444	14 BEVCOMGEN	L04653 Venezuelan
10	2120.6	42.4	11423	14 VEUS5347	US5347 Venezuelan
11	2119.6	42.4	11447	14 EVN8SPERPA	L01442 Venezuelan
12	2119	42.4	11420	14 AF004459	AF004459 Venezuela
13	2119	42.4	11420	14 VEUS5360	US5360 Venezuelan
14	2119	42.4	11423	14 AF375051	AF375051 Venezuela
15	2119	42.4	11423	14 VEUS5342	US5342 Venezuelan
16	2118	42.4	11444	14 EVN8SPENV	J04332 Venezuelan
17	2118	42.4	11445	14 AF069903	AF069903 Venezuela
18	2118	42.4	11446	14 EVN8SPERB	L01443 Venezuelan
19	2118	42.4	11459	6 AR170868	AR170868 Sequence
20	2117.4	42.3	11423	14 VEUS5345	US5345 Venezuelan
21	2117.4	42.3	11423	14 VEUS5350	US5350 Venezuelan
22	2115	42.3	11465	14 AF075257	AF075257 Venezuelan
23	2114.8	42.3	11422	14 EVN8SPERCF	L00930 Venezuelan
24	2114.2	42.3	11421	14 AF100566	AF100566 Venezuelan
25	2103.6	42.1	11391	14 AF075253	AF075253 Venezuelan
26	2087.6	41.8	11464	14 AF448535	AF448535 Venezuela
27	2087.6	41.8	11464	14 AF448538	AF448538 Venezuela
28	2086.4	41.7	11385	14 AF075259	AF075259 Venezuela
29	2086	41.7	11461	14 AF448537	AF448537 Venezuela
30	2085.6	41.7	11344	14 AF075256	AF075256 Venezuela
31	2085.6	41.7	11464	14 AF448536	AF448536 Venezuela
32	2084.8	41.7	11464	14 AF448539	AF448539 Venezuela
33	2081.8	41.6	11395	14 AF075251	AF075251 Venezuela
34	2074.8	41.5	11464	14 VEUS34999	U04999 Venezuelan
35	2070	41.4	11311	14 AF075255	AF075255 Venezuela
36	2049.2	41.0	11468	14 AF075252	AF075252 Venezuela
37	2039.4	40.8	11494	14 AF075258	AF075258 Venezuela
38	2031	40.6	11530	14 AF075254	AF075254 Venezuela
39	1676.8	33.5	11657	14 KRVBNGC	W20162 Rose River
40	1631.8	32.6	11698	14 AB032553	AB032553 Sagiyama
41	1589.2	31.8	14529	14 SFVRNAIS	Z48163 Semliki for
42	1584.8	31.7	7422	14 SF0251359	AJ251359 Semliki F
43	1573.8	31.5	11411	14 AF237947	AF237947 Mayaro vi
44	1567.4	31.3	11488	14 BFU73745	U73745 Barman Fore
45	1567.4	31.3	11822	14 AF079456	AF079456 O'nyong-n

ALIGNMENTS

RESULT 1	AF214040	11484 bp	RNA	linear	VRL 27-JAN-2000
LOCUS	AF214040				
DEFINITION	Western equine encephalomyelitis virus strain 71V-1658				
ACCESSION	AF214040				
VERSION	AF214040				
KEYWORDS	complete cds.				
SOURCE	AF214040.1				
ORGANISM	Western equine encephalomyelitis virus				
REFERENCE	Western equine encephalomyelitis virus				
AUTHORS	Alphavirus; ssRNA positive-strand viruses, no DNA stage; Togaviridae;				
	Necolitzky,D.J., Schmalitz,F.L., Parker,M.D., Rayner,G.A.,				

TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL
Fisher,G.R., Trent,D.W., Bader,D.E. and Nagata,L.P.						Complete genomic RNA sequence of western equine encephalitis virus	
Complete genomic RNA sequence of western equine encephalitis virus						and expression of the structural genes	
J. Gen. Virol. 81 (Pt 1), 151-159 (2000)							
2 (bases 1 to 11484)							
Netolitzky,D.J., Schmaltz,F.L., Parker,M.D., Rayner,G.A.,							
Fisher,G.R., Trent,D.W., Bader,D.E. and Nagata,L.P.							
Direct Submission							
Submitted (08-DEC-1999) Medical Countermeasures Section, Defence							
Research Establishment Sufield, P.O. Box 4000, Stn Main, Medicine							
Hat, Alberta T1A 8K6, Canada							
Location/Qualifiers							
1..11484							
/organism="Western equine encephalomyelitis virus"							
/mol_type="genomic RNA"							
/strain="71V-1658"							
/db_xref="taxon:11039"							
1..11484							
/note="sequence of this entry added to the sequence of							
AF214039, strain CBA87, constitutes the Western equine							
encephalomyelitis virus complete genome"							
<1..24							
25..7428							
/note="encodes nsp1, nsp2, nsp3 and nsp4 proteins"							
/codon_start=1							
/transl_except={pos:5581..5583,aa:OTHER}							
/product="nonstructural polypeptide"							
/protein_id="NP28339.1"							
/db_xref="GI:6760411"							
/translation="MERITVLDADSDPYKSLQRTPEFIEIAQVITDNDHANAPSES							
HVAKLIIESEVDQIVLIDISGAPVRHNSHRHCHICPMISADDPDLQRYARLKAK							
SDIDKNIASKADLLTEWMTSPDAPETSCHTCATQCFSSVAYQVVAHAPTS							
YHOLKGVRTYTWIGDFTPTMYKPMASGYPTVNTMADBEVLEARNGLGNSDIOSES							
RLGSLTIKRRRLQPTNKKIIFSVGQITIEDRLISMLPNVHLKNGSKNRCGQSG							
IYSEEGYIKKTIITSPGIXGVENLASMHBEGFLSCVYDTLIGERSVPTCYVPP							
TLCSQMTGILATIDVSDAQKLVGLNRIYVNGTQNTNTMVOYLIPVAQAFSR							
AREHRADLDKEKLGVERETLTGCGCAAFKQTKITSYKKGQGTIKKVPVAFS							
PRLTSHGLDMGFRRLKLLLEPTVPAPAIIMADVENHRLQGEAEVAAEERREAL							
PLPILEIKETEVEAEVLDMQAGSGVETRGHRIATVSYGEEKIGSAYLSPOAV							
NSEKLACIHLPAEOVLWMTHGGRAGRYKVEPHGKVIYPECTAVPVDPOFALSAT							
VENREPEVNRFLHHIANGALNDBEYKTVKQDPTSEVYPIIDAKCYKRDAG							
LCLQGLDVEPFFHEFAYESLKLTPAPAPKVPFTTNGVGVGSGSGIISATYKQDVA							
SAKENCAGEIIRVDRMRMDVAARVADSVLLNGKVPVNLVIDEAPCAHAGTLLA							
IAIYAPKVVLCGDPKQCGFENMCMCKVHFNDICTEVYHKSISRCTQATVSTIT							
FYDRMKTVPNCADKIIDTTGTTKPHNDILITCFRSMVQLDIDYGNHIMAAAS							
QGLRGKGYAAYRVKYNENPLYSOTSEHVAENLITREKRIWKTLAGDPMITLAKV							
GDPLASIDDMORHEDAIIMARVLDKPDQVADONKYNQWMAELPVALTAIVLTROO							
METIHPFRKHDAYSPEMALNFCFRFVDDISGLFASPYATALTYRDOHMSGKKN							
YGLGVEAKELSSRRYPCITKAVDTGRVADIKNNTKIDYSPITINVPPLMRILPHSLIV							
HKGQGTDEHSGFLSKMKGSVLIQDPIISGKKVESGPIIPVITIRCDLGLPSHV							
GYKDIIFVNVRTPYRNHNYQOCEBHAIHSMMLTCAVHNLGTGCAIVGLADRAPI							
ENIITVAASPRFTRVCPQKNTAENTVELTFEFGDNHNLHTDGLDGLVINDIYOGS							
TVTRAGARAYRAVRIGDIDSKADQAIYNAASQOGSGVGLYKMKPAFDROPRI							
VGTRAKYHEBLLIHANGPNSKMBEPGDKLAAANSTASINYABRTIKTYSPLIS							
TGLTSGGKDRVQMSLHLFTAFDITDADVYITCLDKQWETRIIEAIRKSEVITLDD							
KPVIDIVLRVHNPSSLAGRPQGYVNEGKLYSLGTRHPMKQAIETAHAMPKSEK							
NEOCLITLIGRSMSSIRSKCVSESEKASPHLTPLCLNVMTAERVYRLSARKEQEG							
AVCSGFLPEKRIITGVQKLCQCSKVLFEQGVPAWPKVKEIILTEPPPATTVIC							
EPTYPERIPSPVIRAPAESLILSGGVSPSSASTMSDID							

[illegible]

QY	541	TGGATAGGCTTGGACACGACCCCTTTATGTATCAAAAAATGCGAGGTTCTTACCTACT	600
Db	541	TGGATAGGCTTGGACACGACCCCTTTATGTATCAAAAAATGCGAGGTTCTTACCTACT	600
QY	601	TACATACGAAACTGGGCTGACGAGAGAGATTTGSAAGCACTGTAACATTGGGCTCGGTAC	660
Db	601	TACATACGAACTGGGCTGACGAGAGAGATTTGSAAGCACTGTAACATTGGGCTCGGTAC	660
QY	661	TCGATCTTTCAGAGAGCAGGCGTTTGGAAAACTCTCAATCTTTAGGAAAGAGGCTCCAA	720
Db	661	TCGATCTTTCAGAGAGCAGGCGTTTGGAAAACTCTCAATCTTTAGGAAAGAGGCTCCAA	720
QY	721	CCTACTAATAGATCATATTTCTCGGTGGTTCAACAATCTACAGAAAGATGATCACTG	780
Db	721	CCTACTAATAGATCATATTTCTCGGTGGTTCAACAATCTACAGAAAGATGATCACTG	780
QY	781	TTAGCTAGCTGGCATCTTCCAAAGGTTCCACTTGAAGGAABCTTACTTCAACAGT	840
Db	781	TTAGCTAGCTGGCATCTTCCAAAGGTTCCACTTGAAGGAABCTTACTTCAACAGT	840
QY	841	AGATGTGGGACCATTTGTGACGTGTGAAGGGTACGTATCAAAAAAGATTAACATCAGCCCA	900
Db	841	AGATGTGGGACCATTTGTGACGTGTGAAGGGTACGTATCAAAAAAGATTAACATCAGCCCA	900
QY	901	GGACTATACGTTAAGTTGAGAACTTGGGCGTCAACATGCAATCGGAGGGTTTCTTGAGT	960
Db	901	GGACTATACGTTAAGTTGAGAACTTGGGCGTCAACATGCAATCGGAGGGTTTCTTGAGT	960
QY	961	TGCCAAGTCACGATTAACGTGTGGCGGCGGAGAGGTTTCTTTTGTGTGTATCGTATGTA	1020
Db	961	TGCCAAGTCACGATTAACGTGTGGCGGCGGAGAGGTTTCTTTTGTGTGTATCGTATGTA	1020
QY	1021	CCAGCCACACTTTGGCATTCAGATGACAGAGGATTTCTGGCCAACTGACGTTAGTGGATAC	1080
Db	1021	CCAGCCACACTTTGGCATTCAGATGACAGAGGATTTCTGGCCAACTGACGTTAGTGGATAC	1080
QY	1081	GCACAAAACTATTGGTTGGGCTCAACCAAGAGATTGTCATAGTATGAGACGCAAGA	1140
Db	1081	GCACAAAACTATTGGTTGGGCTCAACCAAGAGATTGTCATAGTATGAGACGCAAGA	1140
QY	1141	AATATTACAAATGCAAACTATCTATTTTACAGTGTGCGCCAGGCGTTTCCAGGTGG	1200
Db	1141	AATATTACAAATGCAAACTATCTATTTTACAGTGTGCGCCAGGCGTTTCCAGGTGG	1200
QY	1201	GGCGCTGAACATCGTGGCCGACTTTGAGACGAGAAAGAACTAAGGGGTGCGGAGCGCACT	1260
Db	1201	GGCGCTGAACATCGTGGCCGACTTTGAGACGAGAAAGAACTAAGGGGTGCGGAGCGCACT	1260
QY	1261	CTTACTATGGGCTGCTGCTGGGCTTTCAGAACCCGAAAAATCATTCCATCTACAGAG	1320
Db	1261	CTTACTATGGGCTGCTGCTGGGCTTTCAGAACCCGAAAAATCATTCCATCTACAGAG	1320
QY	1321	CCTGTATGCAAACTTATGAAAGTACTGCGCTTTTACTCATTTGTGATTCACGC	1380
Db	1321	CCTGTATGCAAACTTATGAAAGTACTGCGCTTTTACTCATTTGTGATTCACGC	1380
QY	1381	CTTACGACGCAACGCGCTGATTTGGGCTTCCGCGTAGGCTTCAAGCTGCTGTAACA	1440
Db	1381	CTTACGACGCAACGCGCTGATTTGGGCTTCCGCGTAGGCTTCAAGCTGCTGTAACA	1440
QY	1441	ACTGTCAAAACCGGCAACGGGCTATTACAAATGGCCGATGTGGAGCATCTGCGTGGTTACG	1500
Db	1441	ACTGTCAAAACCGGCAACGGGCTATTACAAATGGCCGATGTGGAGCATCTGCGTGGTTACG	1500
QY	1501	CAGGAAGCTGAAGAGTGGCTGTGACGCGAAGAGATCAAGAAAGCCCTGCTCAC	1560
Db	1501	CAGGAAGCTGAAGAGTGGCTGTGACGCGAAGAGATCAAGAAAGCCCTGCTCAC	1560
QY	1561	CCTGAATATGAAAAAGAGACCGTTAAGGCGAAGTGAACCTATTATGCAAGAGCGCAGA	1620
Db	1561	CCTGAATATGAAAAAGAGACCGTTAAGGCGAAGTGAACCTATTATGCAAGAGCGCAGA	1620
QY	1621	GCAGGTAGCGTGGAGACACACGAGAGACATCAGGCTGTGACAAAGTTTACCAGCGCAAG	1680

Db	1621	GCAGGTACCGTGGAGACACACGAGGACACATCAGGGGTGACAAAGTTACCCAGCGGAAGAG	1680
Qy	1681	AAGATGGGGCTTACGGCTATACCTTTCACCCCGAGCGGATTTGAATAGGAAAACTGGCG	1740
Db	1681	AAGATGGGGCTTACGGCTATACCTTTCACCCCGAGCGGATTTGAATAGGAAAACTGGCG	1740
Qy	1741	TGTATCCACCCATTTGGCGGAACAAGTACTGTGTAATGACTCACAAAGGTAGGGCGAGAGA	1800
Db	1741	TGTATCCACCCATTTGGCGGAACAAGTACTGTGTAATGACTCACAAAGGTAGGGCGAGAGA	1800
Qy	1801	TACCAAGTCGAGCCATACACGGTAAAGTCAATTGTACCAGAAGGACGGCGGTCCCTGTT	1860
Db	1801	TACCAAGTCGAGCCATACACGGTAAAGTCAATTGTACCAGAAGGACGGCGGTCCCTGTT	1860
Qy	1861	CAAGACTTCAGGCAATTGAGTAAGGAGCTACGATCGTTTTCAACGAGAGGAGATTGGTA	1920
Db	1861	CAAGACTTCAGGCAATTGAGTAAGGAGCTACGATCGTTTTCAACGAGAGGAGATTGGTA	1920
Qy	1921	AACAGATACCTGCACCAATCGCAATCAACGGAAGAGCGGCTAAAACACTGACGAAAGATAC	1980
Db	1921	AACAGATACCTGCACCAATCGCAATCAACGGAAGAGCGGCTAAAACACTGACGAAAGATAC	1980
Qy	1981	TATAAGACTGTAAAGACTCAGAGACACAGACTCAGAATACGTCCTTCGATTTAGACGACGA	2040
Db	1981	TATAAGACTGTAAAGACTCAGAGACACAGACTCAGAATACGTCCTTCGATTTAGACGACGA	2040
Qy	2041	AAGTGTGTTAAGCGAAGAAAGCGAGGTCCTTGTCCTAACCCGGTATCTGTGTAGATCCA	2100
Db	2041	AAGTGTGTTAAGCGAAGAAAGCGAGGTCCTTGTCCTAACCCGGTATCTGTGTAGATCCA	2100
Qy	2101	CCATTTCCAGAGTTTGGGTGCGTACGAGAGTCTCAAGACACGACCGAGAGACCTCACAAAGTC	2160
Db	2101	CCATTTCCAGAGTTTGGGTGCGTACGAGAGTCTCAAGACACGACCGAGAGACCTCACAAAGTC	2160
Qy	2161	CCAAACCATCGGAGTCTATGAGTCCAGGTTCAGGTAAATCTGGAAATCATCAAAAGCGCT	2220
Db	2161	CCAAACCATCGGAGTCTATGAGTCCAGGTTCAGGTAAATCTGGAAATCATCAAAAGCGCT	2220
Qy	2221	GTCGACTAAGAAAGATCTGTGTGTGTAAGTCCGAAGAAAGAAAACTGGCGGAATTCATAGG	2280
Db	2221	GTCGACTAAGAAAGATCTGTGTGTGTAAGTCCGAAGAAAGAAAACTGGCGGAATTCATAGG	2280
Qy	2281	GATTAAGGAGAGATGAGACGTATGAGATGTTGCTCTAGACCTGTCCGATTCAGTCTCTTA	2340
Db	2281	GATTAAGGAGAGATGAGACGTATGAGATGTTGCTCTAGACCTGTCCGATTCAGTCTCTTA	2340
Qy	2341	AATGGGGTTAAGCACCCTGTTAACACTCTGTACATTGATGAGGCAATTTGCTCGCATGCA	2400
Db	2341	AATGGGGTTAAGCACCCTGTTAACACTCTGTACATTGATGAGGCAATTTGCTCGCATGCA	2400
Qy	2401	GGGACGCTGCTGGCACTGATTTGCCATGTCAAACTTAAGAAAGTGTATTGTGCGGGAC	2460
Db	2401	GGGACGCTGCTGGCACTGATTTGCCATGTCAAACTTAAGAAAGTGTATTGTGCGGGAC	2460
Qy	2461	CCAAAACAATCGGCTCTTTTAACATGATGAGCTGTAAGAGTCAATTTTAACATGACATA	2520
Db	2461	CCAAAACAATCGGCTCTTTTAACATGATGAGCTGTAAGAGTCAATTTTAACATGACATA	2520
Qy	2521	TGCACTGAAGTGTACCAATAAAAGCATCTTAGAGAGTGCACACAGACTGTAAACCGCATC	2580
Db	2521	TGCACTGAAGTGTACCAATAAAAGCATCTTAGAGAGTGCACACAGACTGTAAACCGCATC	2580
Qy	2581	GTCCTCCAGCGCTTCTTACGACACGCGAATGAAAGCGGTTAACCATGTGCTGTAAATATC	2640
Db	2581	GTCCTCCAGCGCTTCTTACGACACGCGAATGAAAGCGGTTAACCATGTGCTGTAAATATC	2640
Qy	2641	ATCATAGATACCAAGGAGCAACAAACCGGACAAATATGATCTGATTCACCTGTTC	2700
Db	2641	ATCATAGATACCAAGGAGCAACAAACCGGACAAATATGATCTGATTCACCTGTTC	2700
Qy	2701	AGAGATGGGTGAACAGCTTACAGATTTGACTACAAATACGAAATCATGACTCGCGCT	2760

Dh 2701 AGAGGATGGGTGAAGACAGTACAGATTGACTACAAATAATCATGATGCGGCT 2760
Qy 2761 GCATCGCAAGGACTTACGCGGAAAGGCGTTATGCTGTCAGGTACAAAGTCAAGAGANT 2820
Db 2761 GCATCGCAAGGACTTACGCGGAAAGGCGTTATGCTGTCAGGTACAAAGTCAAGAGANT 2820
Qy 2821 CCACCTACTGCGAGACTTTCAGGACGCTGAACGCTGTTACTTACAGCGACGAAAGAGC 2880
Db 2821 CCACCTACTGCGAGACTTTCAGGACGCTGAACGCTGTTACTTACAGCGACGAAAGAGC 2880
Qy 2881 ATTGTCGGAAGACGCTAGTGTGATCCCTGGAATAAGACACTTACAGCTAAATATCC 2940
Db 2881 ATTGTCGGAAGACGCTAGTGTGATCCCTGGAATAAGACACTTACAGCTAAATATCC 2940
Qy 2941 GGGGATTTCAAGGCTTCAATTGAGACGCTGCGAAGCGGCAACGACGCTATATGCGAC 3000
Db 2941 GGGGATTTCAAGGCTTCAATTGAGACGCTGCGAAGCGGCAACGACGCTATATGCGAC 3000
Qy 3001 GTTCTGTATAGCCGCGACAGCTGATGTGTTCCAGAAATAGTGAACGTCGCTGGGCG 3060
Db 3001 GTTCTGTATAGCCGCGACAGCTGATGTGTTCCAGAAATAGTGAACGTCGCTGGGCG 3060
Qy 3061 AAGGCTTTAGACCAAGTCTTGCCCAAGCAATGTGCTGACGAGACAGCAGTGGAG 3120
Db 3061 AAGGCTTTAGACCAAGTCTTGCCCAAGCAATGTGCTGACGAGACAGCAGTGGAG 3120
Qy 3121 ACGTTGCAACCATTCAGCATGACAGACGCTACTCACTGAAATAGGCACTGACTTT 3180
Db 3121 ACGTTGCAACCATTCAGCATGACAGACGCTACTCACTGAAATAGGCACTGACTTT 3180
Qy 3181 TGCACACAGGTTCTTTGAGTAGACCTGACAGTGGGTTATTTCCGCTCTACCGTCGCA 3240
Db 3181 TGCACACAGGTTCTTTGAGTAGACCTGACAGTGGGTTATTTCCGCTCTACCGTCGCA 3240
Qy 3241 CTTAATTACAGGATCAGCACTGAGATATCTGCCAGGAAAGAACTGATGGGCTTAT 3300
Db 3241 CTTAATTACAGGATCAGCACTGAGATATCTGCCAGGAAAGAACTGATGGGCTTAT 3300
Qy 3301 AGAAGGTAGCAAGAGATTTGTCACGCGGATATCCGTGCATCAAAAACGCTTGACCA 3360
Db 3301 AGAAGGTAGCAAGAGATTTGTCACGCGGATATCCGTGCATCAAAAACGCTTGACCA 3360
Qy 3361 GGCAGGAGTGTATATAGGAATATACCATCAAGACTCTCCCAAAATATATGTG 3420
Db 3361 GGCAGGAGTGTATATAGGAATATACCATCAAGACTCTCCCAAAATATATGTG 3420
Qy 3421 GTTCCATTAAATCGCCGCTGCCCACTGTTGATCGTTGACCAAAAGGACAGGCTACA 3480
Db 3421 GTTCCATTAAATCGCCGCTGCCCACTGTTGATCGTTGACCAAAAGGACAGGCTACA 3480
Qy 3481 ACTGATCAGCGGATTTCTTATCTAAGTGAAGGCAAACTGTGTGTGTATGCGCAT 3540
Db 3481 ACTGATCAGCGGATTTCTTATCTAAGTGAAGGCAAACTGTGTGTGTATGCGCAT 3540
Qy 3541 CCTATCAGCATTTCCAGGAAAGAACTAGATGTCATGCTGCTGCTGCTGCTGCTGCT 3600
Db 3541 CCTATCAGCATTTCCAGGAAAGAACTAGATGTCATGCTGCTGCTGCTGCTGCTGCT 3600
Qy 3601 AGGTGTGATCTGATTTGGAAATACCTAGCCATGTGCTGCTGCTGCTGCTGCTGCT 3660
Db 3601 AGGTGTGATCTGATTTGGAAATACCTAGCCATGTGCTGCTGCTGCTGCTGCTGCT 3660
Qy 3661 AATGTAGGACCCGCTACAGAACATCACTACCAAGTGCAGAGATCAGCTATTCAC 3720
Db 3661 AATGTAGGACCCGCTACAGAACATCACTACCAAGTGCAGAGATCAGCTATTCAC 3720
Qy 3721 CACAGCATGCTAACGTTAAGGCTGCTGCAACCTGGAACAGTGGGAAATGTGCT 3780
Db 3721 CACAGCATGCTAACGTTAAGGCTGCTGCAACCTGGAACAGTGGGAAATGTGCT 3780
Qy 3781 ATAGGATAGGCTTGTGATGCGCAACCGAATATCATCTGCGGTGGCAGCTCA 3840
Db 3781 ATAGGATAGGCTTGTGATGCGCAACCGAATATCATCTGCGGTGGCAGCTCA 3840

Qy 3841 TTTAGGTTTACCCGCTGCTGTCAGGCTTAAGAACACTGCCGAAATATAGTGGTTCTTTC 3900
Db 3841 TTTAGGTTTACCCGCTGCTGTCAGGCTTAAGAACACTGCCGAAATATAGTGGTTCTTTC 3900
Qy 3901 GTGTTCTTCGCAAGGACCAACGCGCAACCAACATAGCAGAGACAGACTCGGTGATG 3960
Db 3901 GTGTTCTTCGCAAGGACCAACGCGCAACCAACATAGCAGAGACAGACTCGGTGATG 3960
Qy 3961 CTTGACAACTCTATCAAGGCTCAACAGTACAGGACAGGAGAGCTCCAGGCTACAGA 4020
Db 3961 CTTGACAACTCTATCAAGGCTCAACAGTACAGGACAGGAGAGCTCCAGGCTACAGA 4020
Qy 4021 GTGATCAGAGTGCATTAGCAAGAGGCTGACCAAGCTATCGTTAATGCTGCTAATAGC 4080
Db 4021 GTGATCAGAGTGCATTAGCAAGAGGCTGACCAAGCTATCGTTAATGCTGCTAATAGC 4080
Qy 4081 AAAGTCAACCAAGTTCGGAAGTGTGCGGTGCACTGTACCGAAATGCGCGCTGCTTTT 4140
Db 4081 AAAGTCAACCAAGTTCGGAAGTGTGCGGTGCACTGTACCGAAATGCGCGCTGCTTTT 4140
Qy 4141 GATAGACGCCAATAGCTGTGCGGACGCGTACACTTGTGAACGACCAACGCTCATATA 4200
Db 4141 GATAGACGCCAATAGCTGTGCGGACGCGTACACTTGTGAACGACCAACGCTCATATA 4200
Qy 4201 CATGCTGAGACCCCAATTTTCTAAGATGCGGAAACCGAAGGCGCACTTAAGCTGCA 4260
Db 4201 CATGCTGAGACCCCAATTTTCTAAGATGCGGAAACCGAAGGCGCACTTAAGCTGCA 4260
Qy 4261 GCTGCTTACATGACATAGCGCTCATGTCACAGCTGAGCGGATTAACAAATATACGTA 4320
Db 4261 GCTGCTTACATGACATAGCGCTCATGTCACAGCTGAGCGGATTAACAAATATACGTA 4320
Qy 4321 CCGCTACTGTCAACCGGCACTTATCTGTGTGCGAAAGATGAGTGAATCATTTGCAT 4380
Db 4321 CCGCTACTGTCAACCGGCACTTATCTGTGTGCGAAAGATGAGTGAATCATTTGCAT 4380
Qy 4381 CACCTGTTCCTGCTTTCGACACCTAGCGGATGCTGCACTATATGCTTGGATATA 4440
Db 4381 CACCTGTTCCTGCTTTCGACACCTAGCGGATGCTGCACTATATGCTTGGATATA 4440
Qy 4441 CAATGGAGACCAAGATATAGAGGCACTTACACCGCAAAAGACGTCGAATTTCTGAT 4500
Db 4441 CAATGGAGACCAAGATATAGAGGCACTTACACCGCAAAAGACGTCGAATTTCTGAT 4500
Qy 4501 GATGACAGCCAGTACATTTGCTGAGGCTCAACCCAAACAGCTCTTTGGCAGGC 4560
Db 4501 GATGACAGCCAGTACATTTGCTGAGGCTCAACCCAAACAGCTCTTTGGCAGGC 4560
Qy 4561 AGACCAAGTTAATCCGTCATAGAGGCAAGTGTATTCATACCTGGAAGGTAACGATTC 4620
Db 4561 AGACCAAGTTAATCCGTCATAGAGGCAAGTGTATTCATACCTGGAAGGTAACGATTC 4620
Qy 4621 CATCAGACCGCAAGACATTTGCGGAATCATAGCAATGTGCGCCAAATCTGAGGCT 4680
Db 4621 CATCAGACCGCAAGACATTTGCGGAATCATAGCAATGTGCGCCAAATCTGAGGCT 4680
Qy 4681 AATGAGCAATTTGCTGTATCTCTGCGGAGATATGTCACAGATCCGCTCCAAATGTC 4740
Db 4681 AATGAGCAATTTGCTGTATCTCTGCGGAGATATGTCACAGATCCGCTCCAAATGTC 4740
Qy 4741 CCAATAGAGGATGAGAGGCGTGTCTCAACCTGACACCTTCATGCTGCTGCTGCTGCT 4800
Db 4741 CCAATAGAGGATGAGAGGCGTGTCTCAACCTGACACCTTCATGCTGCTGCTGCTGCT 4800
Qy 4801 GCTATGACGCGTGAAGCGCTATACAGGTTGCGCTCTGCGAAGAAAGAGTGGCGCTA 4860
Db 4801 GCTATGACGCGTGAAGCGCTATACAGGTTGCGCTCTGCGAAGAAAGAGTGGCGCTA 4860
Qy 4861 TGTCTATCATTTCTGTTGCGAAGTACAGATACAGGCTGCGAAGCTACAGTGCAGC 4920
Db 4861 TGTCTATCATTTCTGTTGCGAAGTACAGATACAGGCTGCGAAGCTACAGTGCAGC 4920

QY 4921 AACACGCTCTGTTTTCAGGCGCTGCTACCAACCGCTGTACACCCAGAGAGTACCGGAA 4980
|||||
Db 4921 AACACGCTCTGTTTTCAGGCGCTGCTACCAACCGCTGTACACCCAGAGAGTACCGGAA 4980
QY 4981 ATAACTTGAAGAACGCCACC 5000
|||||
Db 4981 ATAACTTGAAGAACGCCACC 5000

RESULT 2
LOCUS MEU01065 4471 bp RNA linear VRL 01-JUN-1995
DEFINITION Western equine encephalomyelitis virus nonstructural polyprotein
gene, partial cds.
ACCESSION U01065.1 GI:393033
VERSION nsp1, nsp2, nsp3, nsp4.
KEYWORDS Western equine encephalomyelitis virus
SOURCE Western equine encephalomyelitis virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
Alphavirus; MEV complex.
1 (bases 1 to 4471)
REFERENCE Weaver, S.C., Hagenbaugh, A., Bellet, L., Netesov, S.V., Volchkov, V.E.,
AUTHORS Chang, G.-J., Clarke, D.K., Goussard, L., Scott, T.W., Trent, D.W. and
Holland, J.J.
TITLE A comparison of the nucleotide sequences of eastern and western
equine encephalomyelitis viruses with those of other alphaviruses
and related RNA viruses
JOURNAL Virology 197 (1), 375-390 (1993)
MEDLINE 94025587
PubMed 8105605
2 (bases 1 to 4471)
REFERENCE Weaver, S.C.
AUTHORS Direct Submission
TITLE Submitted (25-AUG-1993) Scott C. Weaver, Biology, University of
JOURNAL California, San Diego, La Jolla, CA 92037-0116 USA
FEATURES
source
1. 4471
/organism="Western equine encephalomyelitis virus"
/mol_type="genomic RNA"
/isolate="5614"
/db_xref="taxon:11039"
/clone="83, 34, 44, 75, 312"
/note="Western equine encephalitis virus"
<1..>4471
/note="Termination codon at position 4253. 4255"
/codon_start=2
/transl_except="(pos:4253..4255,aa:OTHER)
/product="nonstructural polyprotein"
/protein_id="AA67651.1"
/db_xref="GI:393034"
/translation="TIRKVPVAFDSFVPLRITSHQPDIGLFRSRLLLEPTVKAPAL
TMADEVLRGLQGEAEVAAEIREALPPLPEIKETVAEDVLMQAGASVET
PRGRIRVTSYQGEKEIGSVALISPOAVNSKLCIHILAEVLMTHKGAGHYKE
PYHKQIVPEGTAVPVDFOALSSASATVFNREFFVNLHIIANGALMTDEYYK
TVKTQDSEVFPDIDARKCYKREDAGPLCTGLDLPDPHEPAPVESLKTNPAAPIKY
PLTGVGPGSGKGIKSAVTKDILVASKENCAEIRIVRMRBDVAAPVDSV
LNLGKHPVNTLIDBAFCAAGTILALIALVKKPKVLCDDPKCCGFNNMCKLKH
NHDICTEVYHSISRCTQVTAIVSTLFDYKMKTVNPKADKIIDTGTTHKHD
LILCFRGMVQQLIDYKNHEIMTAASQGLRKGVAARVKNVENPLYSOTSHVNV
LVRTTEKRIWKTLAGDPWIKTPTAKYGGDTASLDQSHDAIMARVLDKPTQAV
FONKVNCAKALBPVLATANI VLRQOMETLHPKHDRAVSPMALNFCPTQFQV
LDGLFSAPYALVTRDQHMNSPGKNWYGNLRVAKLSRRVCITKAVDTGAVAD
RNTIKQYSPITNVPLNRLPHSLIDHKQGTDSGFLSKMAGSVLVIGPISI
PKRVESWGPLPTNTRCDLGLPSHVGRKDIIFVNRITYKNHVOQCDHAIIHS
MLTCAVHLNLTGTCVAILGYGLADRAENITIVARFRFRFCOPNTEENELV
VFPRKDNHNDHODRLGVADNPIYOGSTRYEAGRAPYVIRGDISKADQALVNA
NSKQSGSGCGALYRKPAFDRQPIAVGTARLVKHEPLIIHAVGPFSPMBEGD
LKLAAYMSIASIVNAERITKISVLPSTISYSGKORVMOSLHLTPAFTPDADT
IYCDKQWERTITRAIRKESVEIIDDKQVDVLDVAVHNSISLAGPQVSEBAG
YSYLEGRFHOTADIAIHAMMPKSSANBOICLYILGESMSSIRSKCPVEESASA
PPHTPLCCNRYNAETARYLRSAKEQFACSSPPLPKYITGVOKQCKPVLFGS
VPPAVHPRKYAEIILETPSPPTTTVICEPTVERLPSPIVISAESMELISFGVS

Query Match 70.7%; Score 3536; DB 14; Length 4471;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 3594; Conservative 0; Mismatches 75; Indels 3; Gaps 1;
mat_peptide
/product="nsp1 protein"
mat_peptide
/product="nsp2 protein"
mat_peptide
/product="nsp3 protein"
mat_peptide
/product="nsp4 protein"
BASE COUNT 1272 a 1110 c 1116 g 973 t
ORIGIN
FSSATSSSTAMSDYRRFVVTADVHQANTSTWISPSAPGLDVOLPSDDTDSHWSIPS
ALGEVTPSPVODLTAECAPRGLAEIMQDENTAPFOLSHRPVAPARRRPIPSRS
TASAPVPRKPRRTYQOPGVARAISELEDEYIRHONXNYBGA YIFSETGCHL
QOKSROCKLODPIFERBAVHEKTYAPRLDLREKMLQKLOLCASEGRSRYSQ"

QY 1332 AACATTAGAAAGTACCTGCGCTTTGACTCATTTGTGATTCCAGCGCTTACAGGCA 1391
|||||
Db 1 AACATTAGAAAGTACCTGCGCTTTGACTCATTTGTGATTCCAGCGCTTACAGGCA 60
QY 1392 CGGGCTCGATATGGGCTTCCGCGGTAGGCTCAAGCTGCTGTGAACCACTGTCAACC 1451
|||||
Db 61 CGGGCTCATATDAGGGCTTCGCGGTAGGCTCAAGCTGCTGTGAACCACTGTCAACC 120
QY 1452 CGCACCGCTATTACATATGCGCGCATGTGGACATCTGTGGCTTACGCAAGAAAGCTGA 1511
|||||
Db 121 CGCACCGCTATTACATATGCGCGCATGTGGACATCTGTGGCTTACGCAAGAAAGCTGA 180
QY 1512 AGAAGTGTGCTGAGCGGAGAGATCAGAGAAAGCCCTGCACCCCTTGCCTCGTAATGA 1571
|||||
Db 181 AGAAGTGTGCTGAGCGGAGAGATCAGAGAAAGCCCTGCACCCCTTGCCTCGTAATGA 240
QY 1572 AAAAGAGACCGTAGAGGAGAGATAGACCTATTATGCAAGAGGAGGAGGAGTAGCGT 1631
|||||
Db 241 AAAAGAGACCGTAGAGGAGAGATAGACCTATTATGCAAGAGGAGGAGGAGTAGCGT 300
QY 1632 GGAAGACACGAGAGACATCAGGCTGACAAATTACCCAGCGCAAGAAAGTTGGGTC 1691
|||||
Db 301 GGAAGACACGAGAGACATCAGGCTGACAAATTACCCAGCGCAAGAAAGTTGGGTC 360
QY 1692 TTACGCTATCTTACCCCGAGGCGGTATTAATAGTAAACCTGGGTATTCACACC 1751
|||||
Db 361 TTACGCTATCTTACCCCGAGGCGGTATTAATAGTAAACCTGGGTATTCACACC 420
QY 1752 ATGGCGGAGCAAGTACTGTATGACTCACAAAGGTAGGCGAGGAGATCAAAAGTGA 1811
|||||
Db 421 ATGGCGGAGCAAGTACTGTATGACTCACAAAGGTAGGCGAGGAGATCAAAAGTGA 480
QY 1812 GCCATACCAAGTAAAGTCAATGTTACAGAAAGGAGCGGCTCCCTGTTCAAGACTTCA 1871
|||||
Db 481 GCCATACCAAGTAAAGTCAATGTTACAGAAAGGAGCGGCTCCCTGTTCAAGACTTCA 540
QY 1872 GGCATTGAGTGAAGCGCTAGATCGTTTAAAGAGGAGAGTTCGTAACAGTACT 1931
|||||
Db 541 GGCATTGAGTGAAGCGCTAGATCGTTTAAAGAGGAGAGTTCGTAACAGTACT 600
QY 1932 GCACCAATCCCATCAACGAGAGGCGCTAAACACTGACGAGAGTACTTAAGACTGT 1991
|||||
Db 601 GCACCAATCCCATCAACGAGAGGCGCTAAACACTGACGAGAGTACTTAAGACTGT 660
QY 1992 AAAAGCTCAGAGACAGACTAGATAGTCTTGGATTTGACGACGAAAGTGTGTTAA 2051
|||||
Db 661 AAAAGCTCAGAGACAGACTAGATAGTCTTGGATTTGACGACGAAAGTGTGTTAA 720
QY 2052 GCGAAGAGAGCAGAGTCTTGTGCTTAAACCGGATCTGGTATATCCACATTTCA 2111
|||||
Db 721 GCGAAGAGAGCAGAGTCTTGTGCTTAAACCGGATCTGGTATATCCACATTTCA 780
QY 2112 GTTTCGATCAGAGTCTCAAGACAGACGAGCACTTCAAAAGTCCCAACATCGG 2171
|||||

Db 781 GTTTCGCGTAGAGAGTCTCAAGACACGACGACGACCTCACAAGAGTCCCAACATCGG 840
Qy 2172 AGTCTAAGAGAGTCCAGGTTAGGTAATCTCGAATCATCAAAAGCGTGTGACTAAGAA 2231
Db 841 AGTCTAAGAGAGTCCAGGTTAGGTAATCTCGAATCATCAAAAGCGTGTGACTAAGAA 900
Qy 2232 AGATCTGGTTGTGAGTGCAGAGAGAGAGAACTGGCGCAAGAAATCATCAGGAGATGTAAGAG 2291
Db 901 AGATCTGGTTGTGAGTGCAGAGAGAGAGAACTGGCGCAAGAAATCATCAGGAGATGTAAGAG 960
Qy 2292 GATGAGACGTATGAGATGTTGCTGCTAGAGACTGTCGATTCAGTGTCTTAAATGGAGTTAA 2351
Db 961 GATGAGACGTATGAGATGTTGCTGCTAGAGACTGTCGATTCAGTGTCTTAAATGGAGTTAA 1020
Qy 2352 GCACCCCGTTAACTCTGTATCATTTGATGAGGCACTTTCCTCCATGAGAGAGCGTCT 2411
Db 1021 GCACCCCGTTAACTCTGTATCATTTGATGAGGCACTTTCCTCCATGAGAGAGCGTCT 1080
Qy 2412 GGCACGTGATTCGATTCGCAAACTTAAGAAAGTGTATGTGGCGGGAAGCCAAAACAATG 2471
Db 1081 GGCACGTGATTCGATTCGCAAACTTAAGAAAGTGTATGTGGCGGGAAGCCAAAACAATG 1140
Qy 2472 CGGCTTCTTTAATCATGATGCTGCTGAAAGTACATTTTAACCATGACATATGCACTGAAGT 2531
Db 1141 CGGCTTCTTTAATCATGATGCTGCTGAAAGTACATTTTAACCATGACATATGCACTGAAGT 1200
Qy 2532 GTACCATTAAGAGATCTCTAGAGAGTGCACACAGACTGTAAACCGCATCTGCTCAAGCT 2591
Db 1201 GTACCATTAAGAGATCTCTAGAGAGTGCACACAGACTGTAAACCGCATCTGCTCAAGCT 1260
Qy 2592 CTTCTACGACAAAGAGAAATGAAGACGGTTAACCCATGTGCTGATTAATCATATGATATAC 2651
Db 1261 CTTCTACGACAAAGAGAAATGAAGACGGTTAACCCATGTGCTGATTAATCATATGATATAC 1320
Qy 2652 CACAGGACCAAAAGCCGCAAAAGATGTATCTAATCTGTTTCAAGAGATGGGT 2711
Db 1321 CACAGGACCAAAAGCCGCAAAAGATGTATCTAATCTGTTTCAAGAGATGGGT 1380
Qy 2712 GAACACGCTACAGATGATCTCAAAATCATGAAATCTGATGCTGGGTGATGCGCAAG 2771
Db 1381 GAACACGCTACAGATGATCTCAAAATCATGAAATCTGATGCTGGGTGATGCGCAAG 1440
Qy 2772 ACTTACGCGGAAAGCGGTTATGCTGTAGGTAACAAGTCAACGAGAAATCCACTTACTC 2831
Db 1441 ACTTACGCGGAAAGCGGTTATGCTGTAGGTAACAAGTCAACGAGAAATCCACTTACTC 1500
Qy 2832 GCAGACTTCTGAGACGCTGAACGTGTTACTTACACGACAGAAAAACGATTTGTCTGAA 2891
Db 1501 GCAGACTTCTGAGACGCTGAACGTGTTACTTACACGACAGAAAAACGATTTGTCTGAA 1560
Qy 2892 GACGCTAGCTGTGATCCCTGATTAAGACCTTAACGCTTAATATCCCGGGGATTTTAC 2951
Db 1561 GACGCTAGCTGTGATCCCTGATTAAGACCTTAATATCCCGGGGATTTTAC 1620
Qy 2952 GGCTTCAATGAGACGCTGGACGCGGAAACAGACGCGCTTAATGAGACGCTTTGATTA 3011
Db 1621 GGCTTCAATGAGACGCTGGACGCGGAAACAGACGCGCTTAATGAGACGCTTTGATTA 1680
Qy 3012 GCGCGACAGACAGCTGATGTTCCAGAAATAGGTGAACGTCTGCTGGGCGAAAGGCTTTGA 3071
Db 1681 ACCGACAGACAGCTGATGTTCCAGAAATAGGTGAACGTCTGCTGGGCGAAAGGCTTTGA 1740
Qy 3072 GCGCAGTCTTGGCCACGCGCAACATTTGTCTGACGAGACAGACGATGGAGACGTTGACCC 3131
Db 1741 GCGCAGTCTTGGCCACGCGCAACATTTGTCTGACGAGACAGACGATGGAGACGTTGACCC 1800
Qy 3132 ATTGAAGATGACAGAGAGCTACTACGTAATGAGCACTGAATCTTTTGCACAGATT 3191
Db 1801 ATTGAAGATGACAGAGAGCTACTACGTAATGAGCACTGAATCTTTTGCACAGATT 1860
Qy 3192 CTTTGAAGTACGCTGACAGATGGATTATTTTCGCTCTTACCGTGCACCTTAACG 3251
Db 1861 CTTTGAAGTACGCTGACAGATGGATTATTTTCGCTCTTACCGTGCACCTTAACG 1920

Qy 3252 GGATCAGACTGGGATTAATCTGCGAGGGAAGAAATGTATGGGCTTAATAGAGGTAGC 3311
Db 1921 GGATCAGACTGGGATTAATCTGCGAGGGAAGAAATGTATGGGCTTAATAGAGGTAGC 1980
Qy 3312 AAAGAGTGTCAACGCGATATCCGTGATCAACAAAGCGGTTGACACAGGCAAGGTAGC 3371
Db 1981 AAAGAGTGTCAACGCGATATCCGTGATCAACAAAGCGGTTGACACAGGCAAGGTAGC 2040
Qy 3372 TGATATTAAGATATATATCCATCAAGAGACTCTCCACAAATTAATGTGTTCCATTAA 3431
Db 2041 TGATATTAAGATATATATCCATCAAGAGACTCTCCACAAATTAATGTGTTCCATTAA 2100
Qy 3432 TCGCGGGTGGCCCACTCGTTGATTCGTTGACCAAAAGGACAGGGTCAACCTGATCAG 3491
Db 2101 TCGCGGGTGGCCCACTCGTTGATTCGTTGACCAAAAGGACAGGGTCAACCTGATCAG 2160
Qy 3492 CGGATTCCTATCTTAAGATGAAGGCAAAATCTGTGTTGATGATCGGCGATCCTATCAGCAT 3551
Db 2161 TGGAATTCCTATCTTAAGATGAAGGCAAAATCTGTGTTGATGATCGGCGATCCTATCAGCAT 2220
Qy 3552 TCCAGGGAAGAAAGTACAGTTCATGGTCCATTTGCCATTAATCCATCAGGTGTATCT 3611
Db 2221 TCCAGGGAAGAAAGTACAGTTCATGGTCCATTTGCCATTAATCCATCAGGTGTATCT 2280
Qy 3612 CGATTTGGGAATACCTTACGATGCGGTAATATGACATTAATCTTGTCAATGTTAGGAC 3671
Db 2281 CGATTTGGGAATACCTTACGATGCGGTAATATGACATTAATCTTGTCAATGTTAGGAC 2340
Qy 3672 CCCGTACAGGAACCATCATCTACCAACAGTGCAGGATCAGCTATCCACACAGCATGCT 3731
Db 2341 CCCGTACAGGAACCATCATCTACCAACAGTGCAGGATCAGCTATCCACACAGCATGCT 2400
Qy 3732 AACGTGAAGGCTGTCCACCACTGAAACATGCGGAAACATGTGTGCTATAGGTTAG 3791
Db 2401 AACGTGAAGGCTGTCCACCACTGAAACATGCGGAAACATGTGTGCTATAGGTTAG 2460
Qy 3792 GCTTGTGATTCGCGCAACCGGAATATCATCACTGCGGTGGCAAGCTCATTTAGGTTAC 3851
Db 2461 GCTTGTGATTCGCGCAACCGGAATATCATCACTGCGGTGGCAAGCTCATTTAGGTTAC 2520
Qy 3852 CCGTGTCTGTACGCTTAAGAACACTGCGGAAATCTGAGGTTCTCTTGTTCTTCGG 3911
Db 2521 CCGTGTCTGTACGCTTAAGAACACTGCGGAAATCTGAGGTTCTCTTGTTCTTCGG 2580
Qy 3912 CAAGGACAAAGGCAACACACATGACACGAGACACTGGGTATGTGCTTGACAACT 3971
Db 2581 CAAGGACAAAGGCAACACACATGACACGAGACACTGGGTATGTGCTTGACAACT 2640
Qy 3972 CTATCAAGGGTCAACGAGTACGAGGAGGAGGCTCAAGCTGACAGAGTACAGAGG 4031
Db 2641 CTATCAAGGGTCAACGAGTACGAGGAGGAGGCTCAAGCTGACAGAGTACAGAGG 2700
Qy 4032 TGAATTAGCAAGAGCGCTGACCAAGCTATGTTATGCTGTAATAGCAAAAGTCAAC 4091
Db 2701 TGAATTAGCAAGAGCGCTGACCAAGCTATGTTATGCTGTAATAGCAAAAGTCAAC 2760
Qy 4092 AGGTTCCGAGGTGCGGTGACCTGACCGAAATGCGCGCTGCTTTTGTATGACAGCC 4151
Db 2761 AGGTTCCGAGGTGCGGTGACCTGACCGAAATGCGCGCTGCTTTTGTATGACAGCC 2820
Qy 4152 AATAGCTGTGGGAGCGGCTAGACTTGTGAAGACGAACCGCTCATCATATGCTGTAAG 4211
Db 2821 AATAGCTGTGGGAGCGGCTAGACTTGTGAAGACGAACCGCTCATCATATGCTGTAAG 2880
Qy 4212 ACCCAATTTTCTTAAGATGCGGAAACGAGGCGACCTTAAGCTCGAGTGTCTACAT 4271
Db 2881 ACCCAATTTTCTTAAGATGCGGAAACGAGGCGACCTTAAGCTCGAGTGTCTACAT 2940
Qy 4272 GAGCATAGCGTCAATGCTCAACGCTGACGCGATTAACAAAATATCAGTACCGCTACTGTC 4331
Db 2941 GAGCATAGCGTCAATGCTCAACGCTGACGCGATTAACAAAATATCAGTACCGCTACTGTC 3000

OY		4332	AACGGCATCTATTCTGTGGGAAAGATCGAGTATGCATTCATGCACTCCTTTAC	43931
Dd		3001	AAC TGCGACTTA TTTCTGTGGCAAAGATCGAGTATGCATTCATGCACTCCTTTAC	30600
OY		4392	TGCTTTGACA CTACGATGCCGATGTCCACATATATTGCTTGATAAACAATGGGAC	44511
Dd		3061	TGCTTTGACA CTACGATGCCGATGTCCACATATATTGCTTGATAAACAATGGGAC	31200
OY		4452	CAGGATTA TCAGGCACTTTCACCCGAAGAAGCGTCGAATTCTG--GATGATGACA	45080
Dd		3121	CAGGATTA TCAGGCACTTTCACCCGAAGAAGCGTCGAATTCTGATGATGACA	31800
OY		4509	GCCAGTAGACA TTGACTTGTCGAGGGTCACCCCAACAGCTCTTTGGCAGGACGACG	45680
Dd		3181	GCCAGTAGACA TTGACTTGTCGAGGGTCACCCCAACAGCTCTTTGGCAGGACGACG	32400
OY		4569	TTACTCCGTC AATGAGGCAAGTTGTATTCTTACCTGGAAAGTACACGATTCATCAGAC	46280
Dd		3241	TTACTCCGTC AATGAGGCAAGTTGTATTCTTACCTGGAAAGTACACGATTCATCAGAC	33000
OY		4629	CGCCAGGACA TTGCTCCGAATTCATGCAATGTGCCCCAACAAATTTGAGCTAATGACA	46880
Dd		3301	CGCCAGGACA TTGCTCCGAATTCATGCAATGTGCCCCAACAAATTTGAGCTAATGACA	33600
OY		4689	GATTTCCTTA CATTCCGCGGGGGAAGATATGTCAGCATCCGCTCCCAATGCCAGTAGA	47480
Dd		3361	GATTTCCTTA CATTCCGCGGGGGAAGATATGTCAGCATCCGCTCCCAATGCCAGTAGA	34200
OY		4749	GGAGTCAGAG CGTGTGCTCCACCTGCATCAACACTTCATGCTGTGTAATTCAGCTATAC	48080
Dd		3421	GGAGTCAGAG CGTGTGCTCCACCTGCATCAACACTTCATGCTGTGTAATTCAGCTATAC	34800
OY		4809	GGCTGAGCG CGTATACAGTTTGCGCTCTGCGAAGAAAGACAGTTGCCGTAATGCTATC	48680
Dd		3481	GGCTGAGCG CGTATACAGTTTGCGCTCTGCGAAGAAAGACAGTTGCCGTAATGCTATC	35400
OY		4869	ATTCCGTTG CCAGAGTACAGATATCACAGGGGTGGAAGACTACAGTACAGCAACCAGT	49280
Dd		3541	ATTCCGTTG CCAGAGTACAGATATCACAGGGGTGGAAGACTACAGTACAGCAACCAGT	36000
OY		4929	CCTGTTTT CAAGCGTCGTACACCGCGCTGTACACCCCGAAGATACGCGAATAATTTCT	49880
Dd		3601	CCTGTTTT CAAGCGTCGTACACCGCGCTGTACACCCCGAAGATACGCGAATAATTTCT	36600
OY		4989	AGAAAAGCG CACC 5000	
Dd		3661	AGAAAAGCG CACC 3672	
RESULT 3				
WEEVNS				
LOCUS	WEEVNS	4498 bp	RNA	linear VRL 14-FEB-2003
DEFINITION	Western equine encephalomyelitis virus partial gene for nonstructural polypeptide, genomic RNA.			
ACCESSION	X74892			
VERSION	X74892.1	GI:398206		
SOURCE	nonstructural polypeptide; NSP1; NSP2; NSP3; NSP4.			
KEYWORDS	Western equine encephalomyelitis virus			
ORGANISM	Western equine encephalomyelitis virus			
	Virus; ssRNA positive-strand viruses, no DNA stage; Togaviridae; Alphavirus; WEEV complex.			
REFERENCE				
AUTHORS	1 Yufarov,V.P., Volchikov,V.E., Netesov,S.V., Safonov,P.F. and Uryvaev,L.V.			
TITLE	Nucleotide sequence of the 26S mRNA and genes of nonstructural proteins NSP2 and NSP3 of Western Encephalitis virus strain 5614			
JOURNAL	Virologia 28, 86-98 (1992)			
AUTHORS	2 (bases 1 to 4498) Uryvaev,L.V., Volchikov,V.E., Yufarov,V.P., Samokhvalov,E.I., Lebedev,A.I., Safonov,P.F. and Netesov,S.V.			
TITLE	Primary structure of proteins of the nsP2 and nsP3 polymerase complex confirm the recombinant nature of western encephalitis virus			

JOURNAL Dokl. Akad. Nauk SSSR 335, 813-818 (1994)
 REFERENCE 3 (bases 1 to 4498)
 AUTHORS Volichkov, V.E.
 TITLE Direct Submission
 JOURNAL Submitted (01-SEP-1993) V.E. Volichkov, All-Union Institute of Molecular Biology of NPO Vector, Koltsovo, Novosibirsk region, 633159, RUSSIA

FEATURES
 source Location/Qualifiers
 1..4498
 /organism="Western equine encephalomyelitis virus"
 /mol_type="genomic RNA"
 /strain="5614"
 /db_xref="taxon:11039"
 <1..>4498
 /codon_start=2
 /transl_except="(pos:4253..4255,aa:OTHER)
 /product="nonstructural polypeptide"
 /protein_id="CA52868.2"
 /db_xref="GI:28400593"
 /db_xref="GOA:Q89241"
 /db_xref="SPTRMBL:Q89241"
 /translation="TIRKVPVDFSVIPLRTHSPDIGHPRSRLLLEPTVKPAPAH
 TMAVHLRGLQGEAEVAAAEIRALPLPEIKETAEVLDLIMOGAGSVET
 PRGIRVTSYSGEEKISVAILSPQAVLNSKKICHPILAEVIMHNGAGKYE
 PYHKRVYVEBGTANPVQDFQALSASATIVPRERFVNRYLNLHINAGLANTDEYIK
 TVKQIDTSEYFPIIDARKCVKREDAGPLCLTGLVDPPIHEFAYESLKTIRPAHPKV
 PTIGDVGVPSSGSGKIIKSAVTKDILVSAKENCABIIROVRMRMDVARTVDS
 LNLNVEKPVNTLYIDEAPFAGCTLLALIAIVKRVLLCGDPKQCFPMMLCKHNF
 NHDICEVYHKHSISRCTQVTYATVSTLFDYKMKVTNPADKIIIDTGTTRKHOD
 LILCPFGWVKQLQIDYKNEHIMTAASOGITRGVAVYKKNENLYSGTSHVNV
 LVTEKRIYWKTLGADPWIKTFTPAKPGDTTSLDMQSEHDAIMARVLDPKPTADV
 FQNVNVCMALAEPLVATANIPLTRQWETLHFKIDRAVSPMALNFTCTRGVAD
 LDSGLFSAPITVLYRDQHMNSGPKMYGMLNREVAELSRAYCIRKAVDTGFVAD
 RNMTIKDYSPTINVPPLNRPLPHSLIYDHKQGTGDSGFSKNGKSVLIGPIS
 MCKKASGRLPTMTIKTCDLIDGISHVHGKVIDIENVVRPYKNHVOQCEDAIHHS
 LKCEAHMLNTGTCVAIGGLADPAENITIVASFRPTRCQCKRANTENVLYF
 VFESKDNHNTDDRLGVLDNITVQSTRYBARAAVYRIKDISKSAOALVNAN
 NSKQVSGVCGALYRKPAPAFDRQPIAVGTARLVKHEPLIHAVGNFSMPEDRE
 LKTLAAVMSIASVYNARIRIKISVPLSTGIVSGKRVMOSSLHLTAPFTDADY
 IYCDKOMETRIETAIIRKESVELIDDDDPVIDLVRVHNSLSAGPYSVNEGKIL
 YSYEGTRPHQIADIAIEIHAMNPKSEANOICLYLIGSMSTIRSKPSESEASR
 PPHRLPCLCYNAAMERVYLRSAKKROPAVCSFPLPKYRIYVQLQCKPVLFSG
 VPEVAPHPKRYAEIILETPPSPTTIVICEPTVBERIPSVYISABEASLSTFGVS
 FSSASTSSTAMSDYDRFVVTADVHOANTSTWMIIPAPGLVDVLPBDTDSHSIPSS
 ALGEVETSPVQDLTAECARPRGLAEIMQFNTPFQPLSDHREVPAPRRPIPSPS
 TASAPVPKPRRTYQOPGVARAIISAEIDELVYRHSNKRVEAGAVIFSETGCHL
 QOKVRQCKQODPIFERAAVHEKYAPRLDLEREMQLKKQLQCSBGNRSYQSGRPI
 PGLQR"

mat_peptide
 <1..292
 /product="NSP1"
 293..2674
 /product="NSP2"
 mat_peptide
 2675..4273
 /product="NSP3"
 4274..>4498
 /product="NSP4"
 BASE COUNT 1279 a 116 c 1125 g 978 t
 ORIGIN

Query Match 70.7%; Score 3536; DB 14; Length 4498;
 Best Local Similarity 97.9%; Pred. No. 0;
 Matches 3594; Conservative 0; Mismatches 75; Indels 3; Gaps 1;

1332 AACAAATTAAGAAATACCTGCGCGCTTGAATCATTTTGTGATTCACAGCCTTACAGCA 1391
 1 AACAAATTAAGAAATACCTGCGCGCTTGAATCATTTGTGATTCACAGCCTTACAGCA 60

1392 CGGGCTCGATATGAGGCTTCGCGCGTAGGCTCAACAGCTGCTGTGAACCAACTGTCAACC 1451
 61 CGGGCTCGATATGAGGCTTCGCGCGTAGGCTCAACAGCTGCTGTGAACCAACTGTCAACC 120

1452 CGCAACGGCTATTACAATGCGGATGTGGAGCATCTCGTGGCTTACAGCAAGAGCTGA 1511

Db 121 CGCACCGGCTATTCAATGGCCGATGTGAGACACTTGCGTGGCTTACAGCAGAAGCTGA 180
Qy 1512 AGAAGTGGCTGCACGCGGAGAGATCAGAGAAAGCCCTGCGACCTTGGCTCCCTGAATAGA 1571
Db 181 AGAAGTGGCTGCACGCGGAGAGATCAGAGAAAGCCCTGCGACCTTGGCTCCCTGAATAGA 240
Qy 1572 AAAAGAGACCGTAGAGCGAGAGATGACCTTCATTATGACAGAGGAGGAGGAGGATAGCGT 1631
Db 241 AAAAGAGACCGTAGAGGAGAGATGACCTTCATTATGACAGAGGAGGAGGATAGCGT 300
Qy 1632 GGAGACACACGAGAGACATCAGGGTGAACAATTACCAGGCGAAGAGAAAGATTGGCTC 1691
Db 301 GGAGACACACGAGAGACATCAGGGTGAACAATTACCAGGCGGATCCCGGTGAAGAAAGATTGGCTC 360
Qy 1692 TTACGCTATACCTTACCCCGGCGGATTTGATATGAGAAAGAAACCTGGGCTATCCACC 1751
Db 361 TTACGCTATACCTTACCCCGGCGGATTTGATATGAGAAAGAAACCTGAGTATCCACC 420
Qy 1752 ATTGGCGGAACAAAGTACTGTATATGACTCACAAAGTATAGGCGAGAGATCAAAAGTGA 1811
Db 421 ATTGGCGGAACAAAGTACTGTATATGACTCACAAAGTATAGGCGGATCAAAAGTGA 480
Qy 1812 GCCATACCAAGTAAAGTCACTTTGACCAAGAGGAGCGGCGTCCCTGTTCAAGCTTCA 1871
Db 481 GCCATACCAAGTAAAGTCACTTTGACCAAGAGGAGCGGCGTCCCTGTTCAAGCTTCA 540
Qy 1872 GGCAATTGATGAGAGCGCTACGATCGTTTCAAGAGGAGGATTCGTAACAGATACCT 1931
Db 541 GGCAATTGATGAGAGCGCTACGATCGTTTCAAGAGGAGGATTCGTAACAGATACCT 600
Qy 1932 GCACCAATCGCAATCAACGAGAGCGCTAAACACTGACGAGAGATCTATTAAGACTGT 1991
Db 601 GCACCAATCGCAATCAACGAGAGCGCTAAACACTGACGAGAGATCTATTAAGACTGT 660
Qy 1992 AAAAGCTCAGAGACACAGACTAGAAATACGTTGATATGACGACGAAAGTGTGTTAA 2051
Db 661 AAAAGCTCAGAGACACAGACTAGAAATACGTTGATATGACGACGAAAGTGTGTTAA 720
Qy 2052 GCGAGAAAGCGCAGAGTCCCTTGTGCTTACCGGATGTGATGATCCACCATTTCAAG 2111
Db 721 GCGAGAAAGCGCAGAGTCCCTTGTGCTTACCGGATGTGATGATGATCCACCATTTCAAG 780
Qy 2112 GTTTGCGTACGAGAGTCTCAAGACGACGACGACGACCTCAAAAGTCCCAACATCGG 2171
Db 781 GTTTGCGTACGAGAGTCTCAAGACGACGACGACGACCTCAAAAGTCCCAACATCGG 840
Qy 2172 AGCTATGAGAGTCCAGGTTCAAGTAAATCTGGAATCATCAAAAGCGGTGTGACTAAGAA 2231
Db 841 AGCTATGAGAGTCCAGGTTCAAGTAAATCTGGAATCATCAAAAGCGGTGTGACTAAGAA 900
Qy 2232 AGATCTGGTTGTGATGCGAAGAGAAAGAACTGGCGCAAAATCATCAGGAGTGTAAAGAG 2291
Db 901 AGATCTGGTTGTGATGCGAAGAGAAAGAACTGGCGCAAAATCATCAGGAGTGTAAAGAG 960
Qy 2292 GATGAGACGATGATGTTGCTGCTAGAGACTGTCGATTCAGTGGCTTCTAAATGGGGTTAA 2351
Db 961 GATGAGACGATGATGTTGCTGCTAGAGACTGTCGATTCAGTGGCTTCTAAATGGGGTTAA 1020
Qy 2352 GCACCGCGTTAACTCTGTAATGATGAGGCAATTTGCTGCGTCAATGCGAGAGCGTGTCT 2411
Db 1021 GCACCGCGTTAACTCTGTAATGATGAGGCAATTTGCTGCGTCAATGCGAGAGCGTGTCT 1080
Qy 2412 GGCACTGATGCGATTCGTAACCTAAGAAAGTGTATGCGGCGGAGCCCAAAACAATG 2471
Db 1081 GGCACTGATGCGATTCGTAACCTAAGAAAGTGTATGCGGCGGAGCCCAAAACAATG 1140
Qy 2472 CGGCTTCTTTAACTGATGCTCTGAAGTACATTTTAAACATGACATATGCACTGAAGT 2531
Db 1141 CGGCTTCTTTAACTGATGCTCTGAAGTACATTTTAAACATGACATATGCACTGAAGT 1200
Qy 2532 GTACCATTAAGACATCTCTAGAGAGTGCACACAGCTGTAACCGGCATGCTCCACGCT 2591
Db 1201 GTACCATTAAGACATCTCTAGAGAGTGCACACAGCTGTAACCGGCATGCTCCACGCT 1260

Qy 2592 CTTCTACGACAAAGATGAGACGCTTAAACCCATGCTGATTAATAATCATATAGATAC 2651
Db 1261 CTTCTACGACAAAGATGAGACGCTTAAACCCATGCTGATTAATAATCATATAGATAC 1320
Qy 2652 CACAGGACCAAAAGCCGCAAAAGATGATCTGATTTCAACTGTTTCAGAGATGGT 2711
Db 1321 CACAGGACCAAAAGCCGCAAAAGATGATCTGATTTCAACTGTTTCAGAGATGGT 1380
Qy 2712 GAAACGCTACATTTGACTCAAAAAATCAAGAAATCATGACGCGGCTGATCGCAAG 2771
Db 1381 GAAACGCTACATTTGACTCAAAAAATCAAGAAATCATGACGCGCTGATCGCAAG 1440
Qy 2772 ACTTACCGGAAAGCGTTTATGCTGACAGTCAAAAGTCAAGAGATCACTATCTC 2831
Db 1441 ACTTACCGGAAAGCGTTTATGCTGACAGTCAAAAGTCAAGAGATCACTATCTC 1500
Qy 2832 GCAGACTTGTAGACGCTGAACGTTTACTTACACGACAGAAAAACGCAATTGTCTGAA 2891
Db 1501 GCAGACTTGTAGACGCTGAACGTTTACTTACACGACAGAAAAACGCAATTGTCTGAA 1560
Qy 2892 GAGGCTAGCTGATGATCCCTGATTAAGACATTAACGCTAAATATCCGGGGAATTCAC 2951
Db 1561 GAGGCTAGCTGATGATCCCTGATTAAGACATTAACGCTAAATATCCGGGGAATTCAC 1620
Qy 2952 GGCTTCAATTGAGACGCTGAGCGGCAACGACGCAATTAATGAGCAAGCTTTGATTA 3011
Db 1621 GGCTTCAATTGAGACGCTGAGCGGCAACGACGCAATTAATGAGCAAGCTTTGATTA 1680
Qy 3012 GCCGACAGACGCTGATGTTTCCAAATTAAGTGAACGTTGCTGGCGGAAGCTTTAGA 3071
Db 1681 ACCGACAGACGCTGATGTTTCCAAATTAAGTGAACGTTGCTGGCGGAAGCTTTAGA 1740
Qy 3072 GCGAGTGTGCGACCGGCAACGTTGCTGAGAGACACAGTGGGAGAGCTTGCAACC 3131
Db 1741 GCGAGTGTGCGACCGGCAACGTTGCTGAGAGACACAGTGGGAGAGCTTGCAACC 1800
Qy 3132 ATTCAAGCATGACGAGAGCTACTACCTGAATATGSCATCACTTTTTCACAGGTT 3191
Db 1801 ATTCAAGCATGACGAGAGCTACTACCTGAATATGSCATCACTTTTTCACAGGTT 1860
Qy 3192 CTTTGAAGTACCTGACAGTGGGTTATTTTCCGCTCTACCGTCCGACTTACTACAG 3251
Db 1861 CTTTGAAGTACCTGACAGTGGGTTATTTTCCGCTCTACCGTCCGACTTACTACAG 1920
Qy 3252 GATACACACTGGATTAATCCGCAAGGAGAAATATATGAGGCTTAATGAGAGTATG 3311
Db 1921 GATACACACTGGATTAATCCGCAAGGAGAAATATATGAGGCTTAATGAGAGTATG 1980
Qy 3312 AAAAGATGTGACGCGGATTCGTCATCAAAAAGCGGTTGACACAGGAGGATAGC 3371
Db 1981 AAAAGATGTGACGCGGATTCGTCATCAAAAAGCGGTTGACACAGGAGGATAGC 2040
Qy 3372 TGATATTAAGAAATATATCAATCAAGAGACTACTCTCAACAATTAATGTTCCATTAA 3431
Db 2041 TGATATTAAGAAATATATCAATCAAGAGACTACTCTCAACAATTAATGTTCCATTAA 2100
Qy 3432 TCGCGGTTGCGGACCTGATCGTTGACCAAAAGGACAGGATCAACTGATCAAG 3491
Db 2101 TCGCGGTTGCGGACCTGATCGTTGACCAAAAGGACAGGATCAACTGATCAAG 2160
Qy 3492 CGGATTCCTATCTTAAGATGAGAGGCAAAATCTGTTGTTGATGCGGATCTATCAGCAT 3551
Db 2161 TCGATTCCTATCTTAAGATGAGAGGCAAAATCTGTTGTTGATGCGGATCTATCAGCAT 2220
Qy 3552 TCCAGGAGAAAGTATGATCGATCGATTCGATTCGCAATTAATCAATCAGGTGTATCT 3611
Db 2221 TCCAGGAGAAAGTATGATCGATCGATTCGATTCGCAATTAATCAATCAGGTGTATCT 2280
Qy 3612 CGATTTGGGAATCTCTAGAGAGTCCGTTAAATTAATGACATTAATCTTGTCAATGTTAGAC 3671
Db 2281 CGATTTGGGAATCTCTAGAGAGTCCGTTAAATTAATGACATTAATCTTGTCAATGTTAGAC 2340

QY 3672 CCCGTAAGGAAACCATCACTACCAACAGTGGAGATCAAGCTATCCACACAGCATGCT 3731
 Db 2341 CCCATACAAAGAACCATCACTACCAACAGTGGAGATCAAGCTATCCACACAGCATGCT 2400
 QY 3732 AACGTAAGGCTGTCACCACTGTAACATGCGCGGAACATGTGGCTATAGGTATAGG 3791
 Db 2401 AACGTAAGGCTGTCACCACTGTAACATGCGCGGAACATGTGGCTATAGGTATAGG 2460
 QY 3792 GCTTGCTGATGCGCAACCGAATATATCATCTCGGTGGACGCTCATTTAGGTTTAC 3851
 Db 2461 ACTTGCTGATGCGCAACCGAATATATCATCTCGGTGGCTCATTTAGGTTTAC 2520
 QY 3852 CCGTGTCTGTAAGCTTAAGAACCTGCGGAAATACTGAGGTTCTTTGTTTCTTGG 3911
 Db 2521 CCGTGTCTGTAAGCTTAAGAACCTGCGGAAATACTGAGGTTCTTTGTTTCTTGG 2580
 QY 3912 CAAGACAAACGGCAACCAACATGACAGACAGATCTCGGTGATGCTTGGACAAAT 3971
 Db 2581 CAAGACAAACGGCAACCAACATGACAGACAGATCTCGGTGATGCTTGGACAAAT 2640
 QY 3972 CTATCAAGGCTCAACCAAGTACGAGGAGAGAGCTCAGGCTACAGATGATCAGAG 4031
 Db 2641 CTATCAAGGCTCAACCAAGTACGAGGAGAGAGCTCAGGCTACAGATGATCAGAG 2700
 QY 4032 TGACATTAAGCAAGAGCGCTGACCAAGCTATGCTGCTATATAGCAAGTCAACC 4091
 Db 2701 TGACATTAAGCAAGAGCGCTGACCAAGCTATGCTGCTATATAGCAAGTCAACC 2760
 QY 4092 AGGTTCCGAGTGTGCGGTGCACTGTACCGAAATGCGCGGTGCTTTTATAGACAGCC 4151
 Db 2761 AGGTTCCGAGTGTGCGGTGCACTGTACCGAAATGCGCGGTGCTTTTATAGACAGCC 2820
 QY 4152 AATAGCTGCGGAGCGGTAGACTGTGAGACGACACGCGCTCATCATCATGCTGTAGG 4211
 Db 2821 AATAGCTGCGGAGCGGTAGACTGTGAGACGACACGCGCTCATCATCATGCTGTAGG 2880
 QY 4212 ACCCAATTTTCTAAGATGCGGAAACCGAGGCGACCTTAAGCTCGAGCTGCTCAT 4271
 Db 2881 ACCCAATTTTCTAAGATGCGGAAACCGAGGCGACCTTAAGCTCGAGCTGCTCAT 2940
 QY 4272 GAGCATAGCGTTCATGTCACACGCTGAGCGGATTAACAAAATATCATGATCGCTACTGTC 4331
 Db 2941 GAGCATAGCGTTCATGTCACACGCTGAGCGGATTAACAAAATATCATGATCGCTACTGTC 3000
 QY 4332 AACCGGCTATATCTGCTGAGCAAAAGATGAGTGTGATCATCATGCTGCTGCTAC 4391
 Db 3001 AACCGGCTATATCTGCTGAGCAAAAGATGAGTGTGATCATCATGCTGCTGCTAC 3060
 QY 4392 TGCTTTGACACATACGATGCGGATGTCACCATATATGCTTGGATTAACATGGAGAC 4451
 Db 3061 TGCTTTGACACATACGATGCGGATGTCACCATATATGCTTGGATTAACATGGAGAC 3120
 QY 4452 CAGGATTAATGAGGCGCATTCACCGCAAAAGAGCGTGAATTTCTG--GATGATGACAA 4508
 Db 3121 CAGGATTAATGAGGCGCATTCACCGCAAAAGAGCGTGAATTTCTG--GATGATGACAA 3180
 QY 4509 GCCAGTGAACATTAAGTGTGATGAGGTGACCCCAAAAGCTCTTGGAGGACAGACCGAG 4568
 Db 3181 GCCAGTGAACATTAAGTGTGATGAGGTGACCCCAAAAGCTCTTGGAGGACAGACCGAG 3240
 QY 4569 TTATCTCGTCAATGAGGCAAGTTGTATTCATCTGGAAGGTACACAGATTCATCAGAC 4628
 Db 3241 TTATCTCGTCAATGAGGCAAGTTGTATTCATCTGGAAGGTACACAGATTCATCAGAC 3300
 QY 4629 CGCCAAAGACATTCGCGAAATTCATGCAATGTGCGCCAAACAAATCTGAGGCTAATGACCA 4688
 Db 3301 CGCCAAAGACATTCGCGAAATTCATGCAATGTGCGCCAAACAAATCTGAGGCTAATGACCA 3360
 QY 4689 GATTGCTGTTACATCCGCGGAGAGATGATGTCACGATCGCTCCAAATCCCAAGTGA 4748
 Db 3361 GATTGCTGTTACATCCGCGGAGAGATGATGTCACGATCGCTCCAAATCCCAAGTGA 3420
 QY 4749 GAGTCAAGAGCGTGTGCTCCACCTCACACTTCATGCTGCTGTATTAATGAGTATGAC 4808

Db 3421 GAGTCAAGAGCGTGTGCTCCACCTCACACTTCATGCTGTGTAATTAAGTATGAC 3480
 QY 4809 GGCCTGAGGCGGTATACAGGTTGCGCTCGCGAAGAAAGACATTCGCGGTATGCTATC 4868
 Db 3481 GGCCTGAGGCGGTATACAGGTTGCGCTCGCGAAGAAAGACATTCGCGGTATGCTATC 3540
 QY 4869 ATTCTGTTGCGGAAGTACAGGATCACAGGCGGTGACAGAGGTACAGTGCACAAACAGT 4928
 Db 3541 ATTCTGTTGCGGAAGTACAGGATCACAGGCGGTGACAGAGGTACAGTGCACAAACAGT 3600
 QY 4929 CCGTGTTCAGGCGGTGATACACGCGGTGATACACCCGAGAAATACCGGAAATATTCCT 4988
 Db 3601 CCGTGTTCAGGCGGTGATACACGCGGTGATACACCCGAGAAATATTCCT 3660
 QY 4989 AGAAACCGCACCC 5000
 Db 3661 AGAAACCGCACCC 3672

RESULT 4
 LOCUS U01034 11678 bp RNA linear VRL 02-MAR-2000
 DEFINITION Eastern equine encephalomyelitis virus North American antigenic variety nonstructural polyprotein and structural polyprotein genes, complete cds.
 ACCESSION U01034.1 GI:393006
 VERSION nsp1; nsp2; nsp3; nsp4; capsid; E3; E2; 6K; E1.
 KEYWORDS Eastern equine encephalitis virus
 SOURCE Eastern equine encephalitis virus
 ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae; Alphavirus; EEEV complex.
 1 (bases 1 to 11678)
 REFERENCE 1 (bases 1 to 11678)
 AUTHORS Weaver, S.C., Hagenbaugh, A., Beljue, L., Netesov, S.V., Volchkov, V.E., Chang, G.D., Clarke, D.K., Gousses, L., Scott, T.W., Trent, D.W. and Holland, D.J.
 TITLE A comparison of the nucleotide sequences of eastern and western equine encephalomyelitis viruses with those of other alphaviruses and related RNA viruses
 JOURNAL Virology 197 (1), 375-390 (1993)
 MEDLINE 94025587
 PUBMED 8105605
 REFERENCE 2 (bases 1 to 11678)
 AUTHORS Weaver, S.C.
 TITLE Direct Submission
 JOURNAL Submitted (24-AUG-1993) Scott C. Weaver, Biology, University of California, San Diego, La Jolla, CA, 92037-0116
 FEATURES
 source
 1. 11678
 /organism="Eastern equine encephalitis virus"
 /mol_type="genomic RNA"
 /strain="North American antigenic variety"
 /isolate="82V137"
 /db_xref="taxon:11021"
 /clone="pBE14 pBE72 pBE59 pBE700 pBE1900 pBE700"
 1. .46
 47. .7528
 /note="termination codon at position 5681. .5683"
 /codon_start=1
 /transl_except="(pos:5681..5683,aa:OTHER)
 /product="nonstructural polyprotein"
 /protein_id="AAC3734.1"
 /db_xref="GI:393007"
 /translation="MEKVAVDLDADSPVKSLOCPHFIEATQVTDNDHNAARAS
 HLAATKLEGEVDVTDVLDIGSAFVRHTSHKHXCIPKMSAEDPRLRYADKLK
 SDVTDKILASAKADLTFTMSTPDLETSLCMTSTSTRYGSVAIVYDVAVAPTSI
 YVQALKGVTITWIGPDTTPMYNAGAVPTVNTNADSVLEARIYGGSSDLHK
 SPQGVSVTRKKILOPTNKVIFSVGSTIYTERILIRSMHLPNVHLKGTSTGRCT
 IVSCGVVVKITLSPETIGKVDLSTMRBGLSKVDTLRGEVSFPVCTYVA
 TLCDQMTGILATDVSDAQKLVLGNQRLVVGRTQNTNTWQVLLPVQAQFSFC
 REHRADLEDEKGLVRSLSVMGCMVFKTKITSIYKRPQTQIKVPAVFNFSVIP
 QPISYGDIGLRIRIKMLPAKABAPIITEADVHLKQDEDAVAVABAVRALP

mat_peptide
16462
/product="nsp1"
1643.
4024
/product="nsp2"
4025.
5701
/product="nsp3"
5702.
7525
/product="nsp4"
7534.
11678
7592.
11320
/codon_start=1
/product="structural_polyprotein"
/protein_id="AAC53735.1"
/db_xref="GI:393008"
translation="MPPPTLNYPMADINPMAYRDNPPRRRRRPPRPPIAAQIEDL
RSILANITLQKRAPNPAPGAPKAPKRPAPKRPAPATKRRPAPKAPKRPKQKQ
RMCMPLKESKTPPEIMLNGQVNGACVVGAGVKKPLVGRFLDNEQLAIKIKKSAIYD
LEYGQVQOMKSKDTLOYSIDKRPFGFNHAGVYQENNRFLIPRPGVGGKSGSRPILD
NGKRVAYGQVNGESRPLSVYTMNNGKVYVNDPGRGSPMLATVCMLANTYPR
CDQPRMPCETCEKRNHETLMLLEQYDSRADLDLLDAVYCKNAPKTRRDDLTHTYTK
LAPRYADPCNGHSGRCDSPRIAIEEVRDADHAGVIRIQTSAMPGLKTDGVLAMVSP
NGKTSKISIDMLAVRTASPGSLVSHHGYLILAQCPEDTAVTGHGSPKSHHTCTVAAH
KIEPPEVGEIKRHPREHGEVLPCNRHYTHKADQGVYEMHOPGLVADHSLSLHSKAK
VKITVPSGAGQVYVCKCPDVRGEGISLQHTTCDYVCGKRAVILDNKKVYVNSGRLLPR
GEGDTPFKGKLANPVPVAKAKCIATLAPRLVBEKRTILYLNHDHPILYLTTRSGSD
NAPNQMREERTPVNFVYTBGSLAYTMGNHPPKRWAOBSSEGNHGHMPEHVVVYYSD
RPLPTLITGLCTCAIIMVSCVTSWMLFRRNLCTITPYKLAAPNAVQILALACIKIK
PRPADDTQVLVNLMMNNONFPMQTLPLALAILCMMLKCLPCCGAPLALVCGALGALG
AAAYEHNAMPKVGIPYKALVERBGVAPVHLOLVNTRILIPSTNLEYITCKYKTVYK
BSVIVKCGAQTCTSKHPDVOGQVFGVYPRMNGAGACPDNTNTQWSEAVYVNSREE
CSIDBAKAYKHTVQAMVNTITYSVSRADYVYNGETPAKIGDAKILIGPLSSAM
SFPDNKVVNVYGEHVVYDFPEYGTGKASFGDLOSRSTSDLVNTANLTKQRFQAGIG
VHTPTQVAPSGEERWKRDGAPLNDVAPFGSIALERPANCVAGSLPISIDIPDAAG
PRISETPEVSLDECKITECTVAPFGIATVAYASKAGNCIHSPEGVAVYIENNV
TLAEGSSFFHFSSTANIHAPFLQVCTSAVTCCKGQKPRKHIVDYVPAQHTESPAL
SATAMSWIKVLVNGTSAPFLVGLITAVVALVLPFHRH"
7592.
8374
/product="capsid"

mat_peptide	8375..8563	/product="B3"
mat_peptide	8564..9823	/product="B2"
mat_peptide	9824..9994	/product="6K"
mat_peptide	9995..11317	/product="B1"
3'UTR	11321..11678	
BASE COUNT	3314 a 2894 c 2809 g 2661 t	
ORIGIN		
Query Match	55.3%; Score 2766; DB 14; Length 11678;	
Best Local Similarity	72.6%; Pred. No. 0;	
Matches 3605; Conservative	0; Mismatches 1355; Indels 6; Gaps 2	
Oy	14 ATCGATCCAAATGGAAGAAATTCAGTTGACTTATGATGCTGACAGCCGGTATGTCAGT	73
Db	36 ACCATTCGAAATGGAGAAAGTCAATGTTGACTTATGACGACAGACAGCCCATTCGTCAGT	95
Oy	74 GGTACAGCGGACGTTTCCACAAATTGATGATGAAAGCAAGCGAGTCACTGACAAATGACC	133
Db	96 CACTGCAAGATGCTTCCACATTTTGAGATGAAAGCAACGAGGTCACTGCAATGAC	155
Oy	134 ATGCCAATGCCAGACGCTTTCCGATGCGAACCAAGCTCATTTAGAGCGAAGTCGACC	193
Db	156 ATGCTAAATGCTAGGGCGTTTTCGACCTAGCTACTAGCTCATTTAGGGAGAAATGATATA	215
Oy	194 GGGACCAAGTTATCTTTGACATTGGAAATGCGCCCGTCAGACATGCACATTCCAATCAC	253
Db	216 CAGACAGAGTATCCTGATATATGGGAGCGCGCTGTAAAGGACAGACGATTCCAAACTA	275
Oy	254 GCTATCATTTATCTGCCCCATATGATTAAGCGCTGAAGACCCGGACAGACTCAACGGTATG	313
Db	276 AGTACCATGTAATTTGCCCAATGAAGACGCGAGAAACCTTGACAGACTCTACCCCTAG	335
Oy	314 CAGAAAGCTTAAAGAAAGTATGACATTAACGCAAGAAACATAGGCCCTTAAGGGCGGAC	373
Db	336 CAGACAACTTAAAGAAAGATGATGCTACATGACAAATATGATGCTCTAAAGCGCGGAC	395
Oy	374 TGCTGGAAGTCACTGTCACACACGACGAGAGACTCCATCTCTGTGATGACACAGACG	433
Db	396 TGGTAAAGTATGTCGAGCGCTGACGCTGAGACACCTCGTTATGATGACACATGACT	455
Oy	434 CCACGTGATGACTCTTTGGAAGTGAAGATATACCAAGATGTTGACGAGTCCATGAC	493
Db	456 CAACTTGACAGATACACGCGCTCCGGCGCTATATCAGAAATGATATGACAGTGCATGAC	515
Oy	494 CGACATCATCTACCAACAGCGGCTTAAAGATTAAGACAAATTAACATGGAATAGGCTTGG	553
Db	516 GCACCTCAATTAACCAAGGCGCTGAAGGATGACAACTATCTAATGGAATCGGGATTCG	575
Oy	554 ACAGACCCCTTTTATGTAACAAAACATGCGAGGTTCCTACCCCTACTTACACACGAAT	613
Db	576 ATACTACACCGTTCATGTTAAGAACATGGCAGGGCCTTACCTTACATTAACACAAAT	635
Oy	614 GGGCTGACGAGAGATATGGAAGCACGTAACTTGGCCTCGTAACTCAGTCTTCAG	673
Db	636 GGGCGGATGAAGATGTTGGAAAGCAGAAATATAGGGGTGGATGTTCAAGCTTGACG	695
Oy	674 AGAGCAGCGCTGGAAAACCTCTCAATCTCTTGAAGAAGAGGCTCCAACTACTAATAGA	733
Db	696 AAAAAGATTCGAAAGATATCCATTAAGAGAAAGAAATTAACAACCAACCAATTAAG	755
Oy	734 TCATATTCGGTGGTCAACAATCTTACAGAAATATGATCACTGTTAGTAGTAGTGCG	793
Db	756 TAAATATTTCTGGGGGTCAACTATTTATCTGAAGAGAAATATCTGTTAGCCAGTTGGC	815
Oy	794 ATCTTCCAAACGTGTTCCACTTGAAGAAAGTCTTAATCTTCCAGAGTATGATGGAGCA	853
Db	816 ATCTACCTAATGTTTTTCACTTAAAGGTAACATGCTTTTACAGGACAGATGTAACACCA	875
Oy	854 TTGTGAGCTGAGAGGTAAGCTCATCAAAAAGATACGATCAGCCGACGATTAACGGTA	913

Db 876 TCGAGAGCTGCGAAGGTTACGTTGTCAAGAAATTAACGCTCAGTCTGGGATTTACGGGA 935
Qy 914 AAGTTGAGAACTTGGCGTCCACATGCAATCGGAGGTTTCTTGAGTTGCCAAAGTCACAG 973
Db 936 AAGTGAATATCTTCTTCCAGCATGCAACCGAGAGGATTTTAAAGTTGCAAGGTTACG 995
Qy 974 ATACCTGCGCGCGAGAGGTTTCTTGTGTGTGTACGTATGTATGACAGCACACTT 1033
Db 996 AACCTTAAAGAGGAGAGGAGGCTCTTTTCCGTGTATGATGCTGCGCAGGACACTGT 1055
Qy 1034 GCGATCAGATGACAGGATTTCTGGCAACTGACGTTAGTGTGATGACGCAAAAACTAT 1093
Db 1056 GCGACAGATGACCGGGATATCTGGCACTGACGTGATGTGATGACGCGCAGAAAGCTGC 1115
Qy 1094 TGGTTGGGCTCAACCAAGGATTTGTCTCAATGTATGAGACGCAAAATTAACACA 1153
Db 1116 TGGTTGGGCTCAACCAAGGATTTGTCTCAATGTGAGCAACACGTAACCAAAATACCA 1175
Qy 1154 TGCAGAACTATCTATTACAGAGTGTGCGCCAGCGCTTTCAGGTGGCGCGTGAACATC 1213
Db 1176 TGCAGAAATTAATCTATTACAGAGTGTGCGCCAGCGCTTCTCGCGGTG---CGGGAAACC 1232
Qy 1214 GTGCCGACTTGGACGACGAGAAAGAACTAGGGGTGCGGAGCGCACTTTACTATGGCT 1273
Db 1233 GCGCAGACTGTGAGAGCGAAAGGGGTAGGGGTACGGGACGTTCCCTAGTCATGGCT 1292
Qy 1274 GGTGTGGGCTTTCAGAACCCGAAATACATCCATCTCAAGAAAGCTGTACGCAAA 1333
Db 1293 GCTGTGGGCTTTCAGAACTCAAGATCACTCAATTTCAAGAGACCTGTGGACTCAAA 1352
Qy 1334 CAATTAAGAAAGTACTGCGCTCTTGACTCATTTGTGATTCACGCGCTTACAGCCAG 1393
Db 1353 CTATTAAAGAGTCCCGCCGCTTATTCATTTGTCTATCCCAACAACCAACAGCTATG 1412
Qy 1394 GCGTGTATATGAGCTTCCGCGCTAGGCTCAAGCTGCTTGAACCACTGTCAAAACCG 1453
Db 1413 GCGTGTATATGAGTGTGCGTCCGCAATTAAGATGATTCGACCGAAAGAGCACCG 1472
Qy 1454 CACCGGCTATTCAATGCGCCATGTGTGAGCATCTGCGTGGCTTACAGCAAGAGCTGAAG 1513
Db 1473 CCGCAATTAATTACTGAGCGCGAGTGTGCAACCTTAAAGGCTGCAAGATGAGCTGAAG 1532
Qy 1514 AAGTGTGTCAGCGGAGAGATCAGAGAAAGCCCTGCAACCTTGTCCCTGAATAGAA 1573
Db 1533 CCGTGTGTCAGCGCTAAAGCCGTAAGCTGTGACGACCTTCTCACTTCTGCGGAGGTGACA 1592
Qy 1574 AAGAGACGCTAGAGCAGAAAGTGAACCTCATTTATGCAAGAGCGAGAGAGGTAGCTGG 1633
Db 1593 AGGAGACGCTAGAGCTGACATCGACTGATCATGAGAGGCGAGAGGACAGCGTGG 1652
Qy 1634 AGACACACAGAGGACATCAGGCTGACAGATTACCCAGGCGAAGAGAGATTTGGTCTT 1693
Db 1653 AGACACCTTAGACAGACATCAAGGTCACGACGTACCCAGSAGAAATAATGATCGCTGT 1712
Qy 1694 AGCTATATCTTCAACCCAGGCGGTATTGAATAGTGAATAAATGCGGTATCCACCAT 1753
Db 1713 AGCGATGCTTTTCAACAAGCGGTCTTAAACAGGAGAGAGTCTGCTGTATTCACCGT 1772
Qy 1754 TGGCGGAAACAAGTACTGTATGACTCAAAAGTAGGGCAGAGAGATCAAAAGTCAGAC 1813
Db 1773 TAGCTGACAGAGTGTGCTGATGATCAACAAAGGCGCGCAGAGATCAAGAGTAGAGC 1832
Qy 1814 CATACACGCTAGAGTCAATTGTACCAAGAGGAGCGCGGTCTCTGTTCAAGACTTCCAG 1873
Db 1833 CATACACGCTAGAGTCAATTGTCTCTAGGTACAGCTATACCAATCCCGGATTTCCAG 1892
Qy 1874 CATTAGTGAGAGCGCTAGATGTTTTCAAGAGAGGAGTTCGTAACAGATACCTGCG 1933
Db 1893 CTCTAGTGAAGTGCACATAGTATTTAAAGAACGGAGTTCGTTAACCTTACCTTAC 1952
Qy 1934 ACCACATGCGCAATCAAGAGAGGCGCTAAACACTGACGAAAGTACTAATAGACTGTA 1993

Db 1953 ACCACATTCGCTTAAACGAGAGGCGCATTAATACAGATGAAGACTACTAACAAGTTGTGA 2012
Qy 1994 AGACTCAGAGACAGACTCAGAAATACGCTTCTGATTAATGACGACAGAAAGTGTAAAGC 2053
Db 2013 AAAGCATGAGACAGACTCTGATGATGATTTGATGATGACGCAAGAAAGTGTGAAGA 2072
Qy 2054 GAGAAAGCAGAGTCCCTTGTGCTTAACCGGTGATCTGTATGATCCACATTTTCAGAGT 2113
Db 2073 AAGGGAGTCCGGAGACCAATGTGCTGTGTGCGGAGACTAGTTGACCCGCACTTCCAGAT 2132
Qy 2114 TTGCTAGAGAGTCAAGACAGACGACGACCTTCAAAAGTCCCAACCATCGAG 2173
Db 2133 TTGCTAGAGAGTCAAGACAGCTCTGCTGACCAACAAAGTGTCTTACATCGAG 2192
Qy 2174 TCTATGAGTCCAGGTTCAAGTAAATCTGAAATCATCAAAAGCGCTGTACTAAGAAAG 2233
Db 2193 TTTATGAGTCCAGGTTCCGAAAGTGTGATTAATCAAAAGCGCTGTACCAAGCGTG 2252
Qy 2234 ATCTGTTGATGATGCGAAGAAAGAACTGCGCAGAAATCATCAGGATGTAAGAGGA 2293
Db 2253 ATCTGTTGATGATGCGAAGAAAGAAATTTGATGAAATCATTAAGACGTCAAAAGTGA 2312
Qy 2294 TGAAGCTATGATGATGCTGCTAGAGACTGTGATTCAGTGTCTTAAATGGGTTAAGC 2353
Db 2313 TGCGCGCATGAGACATCGCCCGCAGAGATGATTCGCTTAAATGGGTTAAGC 2372
Qy 2354 ACCCGTTAAGCTCTGATCAATGATGAGGATTTGCTGCAATCAGAGAGCGCTGTG 2413
Db 2373 ACTCGTGACACATGATCAATGAGGAGATTCGCTTGCATGACAGGAGCGCTGTAG 2432
Qy 2414 CACTGATGCGATGCTCAAACTTAAAGAGTGTATTTGCGGGGAGCCAAACCAATGCG 2473
Db 2433 CACTTATGCGATGCTCAAGCCAAAGAAATTTGATGATGATGATGATGATGATGATG 2492
Qy 2474 GCTTCTTAAAGATGATGCTGAAAGTCAATTTTAAACATGACATATGACTGAAGTGT 2533
Db 2493 GCTTCTTAAAGTATGATGCTTAAAGTCAATTTTAAACAGATATGACAGAGTGT 2552
Qy 2534 ACCATTAAGATCTCTAGAGGTGACACAGACTGTAAACCGCATGTGTTCACAGCTGT 2593
Db 2553 ATCAAGAGATTTCTGCGGAGTGCATTAAGACGTACATCACTTTTCCACCTGT 2612
Qy 2594 TCTACAGAGAGATGAAGACGTTAACCCATGCTGATTAATTCATATCATATGATCA 2653
Db 2613 TCTATGATTAACGATGAGAACTGTCAACCATGATGATGATGATGATGATGATGATG 2672
Qy 2654 CAGGACCAAAAGCCGACAAAGATGATCTGATTTCTAACCTGTTTCAAGAGTGGTGA 2713
Db 2673 CAGATTAACCAACCTTAAAGATGATTAATTTAACTGCTTTAGAGGTTGGTGA 2732
Qy 2714 AACAGCTACAGATTGATCAAAATCAAGAAATCATGCTGCGGCTGATGCAAGAC 2773
Db 2733 AACAGCTGATGATGATCAAAAGCAAGAGATCATGACCGCAGCGCTTCAAGGAGC 2792
Qy 2774 TTAAGCGGAAAGGCTTATGCTGTCAAGTCAAAAGTCAAGAGATCACTCTGCTGCG 2833
Db 2793 TTAAGCGGAAAGGCTTATGCTGTCAAGTCAAAAGTCAAGAGATCACTCTGCTGCG 2852
Qy 2834 AGACTTCTGACAGTGAACGTTTACTTAAACGACGCAAGAAAGCAATTTCTTGAAGA 2893
Db 2853 AGACATCAGAGCATGTGAATGATTAATTAACCGCATGAAAGATGATGATGAGAA 2912
Qy 2894 CGCTAGCTGTGATCTCTGATTAAGACATTAACGCTTAATTTCCGGGATTTTCAGG 2953
Db 2913 CTTGGCGGAGACCTTGTGATCAAGAGTGAACGATGATTCAGATTAATTTTCAGCG 2972
Qy 2954 CTTGATGAGAGCTGAGCGGACGACGACGACGACGACGACGACGACGACGACGACGAC 3013
Db 2973 CCACACTGAGAAATGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3032
Qy 3014 CGCAGACAGCTATGATGTTCCAGAAATGAAGTGAAGTGTGCTGCGGAGAGCTTTAGAGC 3073
Db 3033 CAGCTAGAGAGAGCTTTTCAAAATTAAGAGTGTGCTGCGGAGAGCTTTAGAGC 3092

QY 3074 CAGTCTGGGCGGCGCAACATTTGCTGAGAGACAGAGTGGGAGAGCTTGACCCAT 3133
 Db 3093 CTGTGTTGGCCACCGCCCATATTTAGCTGACCCGCTCGCATGGGAGCTATTCAGCGT 3152
 QY 3134 TCAGCATGACAGACGCTACTCACCCTGAATATGGAACCTGCTTTTGGACAGAGTTCT 3193
 Db 3153 TCAGAGATGACAAAGCGTATTCGCTGAGATGGCTTAACTTTTTCGACAGATTCT 3212
 QY 3194 TTGAGATGACCTGAGACAGTGGTTATTTCCGCTCTACCGTGCACCTTACACAGG 3253
 Db 3213 TTGGCTGACATGACAGCGGGTGTCTCCGCGCCAACTGTTCCGCTGACTTACACA 3272
 QY 3254 ATCAGACTGGGATTAATCCGCGAGGAGAAACATGATGGCTTAATAGAGAGGTACAA 3313
 Db 3273 ATGAACTGGGATTAATGCCAGGTCCAAACATGATGTTTGTGCACTGGCACTGCTA 3332
 QY 3314 AGGAGTTGTCACGGCGATATCCGTGATCACAAGAAAGCGTTGACACAGCGAGGTAGCTG 3373
 Db 3333 AAGAACTTGACGTCGGTATCTTGTATTTGAAAAGCGTGATACAGGTAGAGTGCTG 3392
 QY 3374 ATATAAGGATTAATACCTCAAGACTACTCCAACTTAATGTGTTCCATTAAATC 3433
 Db 3393 ACGTTCGACAGACATATCAAGACTATTAACCGCTAATAAATGTGTACCCCTGATA 3452
 QY 3434 GCCGATTGCCCATCTCGTTGATCGTTGACCAAGAGACAGGTTACAACTGATCAGACGG 3493
 Db 3453 GAAGACTCCCACTCATTTGTTGTGACACATAGATACCTGGAAACGTTATTAATCTCC 3512
 QY 3494 GATTCTATCTAATGATGAAGGCAAAATCTGTGTGTGTGATCGCGATCTTACAGACTC 3553
 Db 3513 AGCTAGTACCAAGATGACCGGAAAACGTAAGTAGTGAGTAACTATGAACATAC 3572
 QY 3554 CAGGAAAGAAAGTGAAGTCCATGGTCCATTGGCCCACTAAATCAGTCAAGTGTGATCG 3613
 Db 3573 CAGAAAGAGTGAAGCACTAGGCCCAACCCCAATGATATATMAACGGAACCTGG 3632
 QY 3614 ATTGGGAATACCTAGCATGCGTAAATATGACATATCTTTGCAATTTAGGACCC 3673
 Db 3633 ACTGGGATCTCTGCGCTTTAGGCATATGACATATTTTATTAACGTAGAGATC 3692
 QY 3674 CGTACAGGAACCATCACTACCAAGTGCAGAGTCAAGCTATCCACCAAGACATGCTAA 3733
 Db 3693 CTTACCGGACCACTATTAACCAAGTGCAGAGCAATGCGATCCACCAAGACATGCTTA 3752
 QY 3734 CGTTAAGCTGTCACTCACTGAACACTGCGGAACATGTGTGCTATAGGCTATGGGC 3793
 Db 3753 CCAAGAAAGCAGTGAACATTTGAACAAAGCGGTACGTGATCGCATTTGGGCTATGGGA 3812
 QY 3794 TTGCTGATCGGCAACGAGAAATATCACTGCGGTGGCAGCGCTAATTTAGTTTACC 3853
 Db 3813 CTGCGGACAGGCGGAGAACTATCTCGGAGTGGCCCGCTCATTCAGGTTCTAC 3872
 QY 3854 GTGTCTGTCACTTAAGAACTGCCAAATACTGAGTTCTTTCGTGTTCTTCGCA 3913
 Db 3873 GTGTGTGCGACCGAAGTGTGCTGGGAAACACTGAGGTGCGGTTGTTTTCCGCA 3932
 QY 3914 AGGCAACGCGAACCAACATGACAGGACAGACTCGGTGTGTGTGTTTAAACAATCT 3973
 Db 3933 AGGCAACGCGAACCATCTCCAAATCAAGATAGAGTGTGTGTTTAAACAATCT 3992
 QY 3974 ATCAAGGTCACACAGGTAGAGGAGGAGAGCTCCAGCTACAGAGTATCAGAGTG 4033
 Db 3993 ACCAAGGTCACATCAACATAGTGCAGAGACCTGCAATATAGAGTGTGCGGCG 4052
 QY 4034 ACATTAGCAGAGGCGTACCAAGCTATCTTAAATGTCTTAATAGCAAAAGTCAACCG 4093
 Db 4053 ACATTAACAAAGAGCAATGATAGTTATGTTTAAACGCGGCAACAAGGCAACCG 4112
 QY 4094 GTTCCGAGTGTGCGGTGCACTGTAACGAAATGGCGGCTGCTTTGATGACAGCCAA 4153
 Db 4113 GTGCGGTGTGTGTGTCCTTTTACAGGAAGTGGCTGAGCTTTTGTATGAAGACCGG 4172

QY 4154 TAGTGTGCGGAGCGCTAGACTTGTGAAGACGAAACCGCTCATCATATCATGCTGTAGAC 4213
 Db 4173 TAGCACTGGTAAAGCCACCTTGTCTCAAGCATTTCTCCGAACGTATCATGCGTTGGCC 4232
 QY 4214 CCAATTTTTCTTAAGATCCGGAACCGGAGGCGCACTTAAAGCTGCGACTGCTTACATGA 4273
 Db 4233 CTAATTTTCTCGGCTATCAAGAAAACGAGAGACCGAAATTTGTGGAAGTATACATGG 4292
 QY 4274 GCATAGGTCATCTGCAACGCTGAGCGGATTCAAAATAATAGTATCCGCTACTGTCAA 4333
 Db 4293 ACATTCGCAAAATTAATCAACGAGAGTTTACTAAAGTCTCAATTCGTTTATCTA 4352
 QY 4334 CCGGCATCTAATCTGTGAGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 4393
 Db 4353 CCGGCATCTAAGCAGAGTGTGATGAGACAGGTTTATGCAATGCTGAACCATCTATTTACAG 4412
 QY 4394 CTTTGCACCTACGAGATCCGATGTCACATATATGCTTGTGATTAACAAATGGAGACCA 4453
 Db 4413 CCATGATTAATACGAGCAGACATTAATTTTACTGTCTATGATTAACAAATGGAGTCAA 4472
 QY 4454 GGAATATCGAGGCGCATTCACCGAAGAAAGCGTCCAA---ATTCTGATGATGACAGAC 4510
 Db 4473 GAATTAAGAGCTATCATCTGGAAGAAAGCTTGAAGACTTACTGAGAGTACAGAC 4532
 QY 4511 CAGTACATTTGACTTGTGAGGCTCCACCACCAACAGCTCTTTTGGCAGGACGACAGGTT 4570
 Db 4533 CAGTGTACATTTGAATCTGTAGGAGTGCACCGTTTGAAGACTTGTGAGAGTACAGGTT 4592
 QY 4571 ACTCGTAAATGAGGCAAGTTGATCATATCTGGAAGTACAGATTCATACAGCCG 4630
 Db 4593 ATTCACACCCGAGGCGAAGGTATTCGTAAGGAGGAGTGAAGTTCATCAATCACTG 4652
 QY 4631 CCAAGCACTTCCGGAATTCATGCAATGAGGCGCAACAAATGAGGCTTAATGAGACA 4690
 Db 4653 CCAAGCACTTCCGGAATTTAGCTATGAGGCTTAACAAAGAAAGCAACAGGACA 4712
 QY 4691 TTGCTGTATCATCTGAGGAGAGATATGTCAGACATCCGCTCCAAATGCGCAGTAGAGG 4750
 Db 4713 TTTCCTTATACGTGTGAGAGAGATGAAACAGCATCCGCTCATAGTGTCCAGTTGAAG 4772
 QY 4751 AGTCAGGCGTGTCTCACTTCAACACTTCCATGCTGTGTATTAATGCTATGACGG 4810
 Db 4773 AGTCGAGGCGCTCTTCCCTTCAACACTTCCATGCTGTGTATTAATGCTATGACGG 4832
 QY 4811 CTGAGCGCGTATCAAGTTGCGCTCTGCGGAAGAAAGAGTTCGCGGTATGCTATCAT 4870
 Db 4833 CAGGCGAGTTTACAGATTAAGTATGCAAGAAATGAAATTCGAGTTGTTCGCT 4892
 QY 4871 TCTGTTGCCGAAGTACAGATGACAGGCGTGCAGAACTACAGTGCAGCAACAGTCC 4930
 Db 4893 TTCACTTACCGAAATACAGATTAACAGGAGTTTCAAGAAATTCATAGCAATTAACCTGTA 4952
 QY 4931 TGTTCAGGCGTGTACCAACCGGCTGTACACCCCGAAGATGACGC 4976
 Db 4953 TATTCCTGCGACATGATCCCGCGCATATCATCAAGAAATTCGCG 4998

RESULT 5
 EEBVIRNA
 LOCUS 11675 bp RNA linear VRL 08-JUL-1992
 DEFINITION Eastern Equine Encephalomyelitis virus RNA sequence.
 ACCESSION X63135 X67111
 VERSION X63135.1 GI:59185
 KEYWORDS 6K protein; C protein; E1; E2; E3; NS1; NS2; NS3; NS4.
 SOURCE Eastern equine encephalitis virus
 ORGANISM Eastern equine encephalitis virus
 Viruses; ssRNA positive-strand viruses; no DNA stage; Togaviridae;
 Alphavirus; EEEV complex.
 1 (bases 1 to 11675)
 REFERENCE
 1 Volchkov, V.E., Volchkova, V.A. and Neteov, S.V.
 Complete nucleotide sequence of the Eastern equine
 encephalomyelitis virus genome
 Mol. Gen. Microbiol. Virusol. 5, 8-15 (1991)

JOURNAL

MEDLINE 91375524
PUBMED 1896061
REFERENCE 2 (bases 1 to 11675)
AUTHORS Volchokov V.E.
TITLE Direct Submission
JOURNAL Submitted (22-NOV-1991) V.E. Volchokov, All Union Institute of Molecular Biology, Koltoovo, Novosibirsk Region, 630359, USSR
FEATURES
source
1. 11675
/organism="Eastern equine encephalitis virus"
/mol_type="genomic RNA"
/strain="ssp. North American variant"
/db_xref="taxon:11021"
/cell_line="virus grown on vero cells"
/clone_lib="cDNA plasmid clones grown in E.coli JC-5183"
47. 5683
/note="unnamed protein product"
/codon_start=1
/protein_id="CAA44846.1"
/db_xref="GI:59187"
/translation="MEKVHVDLSDSPFVSKLQRCFPEFEIETATQVTDNDANARAS
HLATKLGEVDITDOVILIDISAPVRHSHKTHCICPMKSAEPDLVYADLTK
SDVTDKCIASAAADLITWSTPDAETPSLCMTDSTCRHSVAYQDVAVHAPT1
VYALKGVRITWIGFTDTPMYKMAAGAYPTYNMNADESVLEARNIGLSDLHEK
SPGKYSIMRKKKLOPTNKVIFSVGSTIYEEIRLLRSMHLNPNVFLKGTSGTCRNT
IVSCGEYVVKKITLSPGIVGVNDLASMHEGFLSCVDTDLGERYSFVCTVYPA
TLCDMGITLADTVSDVDAOKLNGLNORITVNGRTONTNTMNYLLPYVAQAFRC
REYRADLEDEKGLGVERSLVWGCCWAKTKITSTYRPGTQITKQPAVENSFVLP
QPTVGPDIGLRRIKMLFDAKKAAPITADVAHLKLGDEAEVAEVAEVAALP
PLPEBVDEIYEADIDLMOBAGASVETPRRHIVITVPEEMIGSAVLEPAVLN
SEKLACHPLAEVULMTHKGRAGVKEPRHRIKVITVPSGTAIPIDPOALSEATIV
FNEBEFVRVYLIHIAVNGALNTDEEYKVVSTETDESEVVIDAKKCVGRGAPV
CLVEIDNPPEHFAVESIKTRPAPAPKVPITGVYGVSGSGKSIKSAVTKRPLVYS
AKKNCHEITIDVCRMGMIDLAATVDSVLNGVHSYDITLIDBAFCHAGTLLALI
AIVPKRVLCGDPKQCGCFNMMLKVHFNHICTEVYHKSISRCTKTVSIVSTLF
YDKMRITVNPONDKIIDTSTYKPKDIIITCFRGWVKQIQUIDYKHEIMTAASQ
GFTRRGVAVYKVENEDVLAQTSSEHVVLLTRTEKRIVMKTLAGDPWIKTLTASYPG
NETRLBEMOEHDAIMAKIILETPASSDVFNQKVCWAKALEPILATANTITRSOW
ETIPAFKODKAVSEPMALNFCRTPFGVIDISGLPSAFTPLASTNEMWDSPPBNKY
GLCMRKAELARRPCLIKAVDTGRVVDRTDITKDINPLINVPFLNRHSHLSVITQ
RYTNGDSQVLTKMTKTVLAVGTGPMNIPKRVETLQSQCTYKALDGIIPALG
KYDII FINVTRPYRHHYQOCEDEHAIHSMILTRAXVDNLNGGTCIALGVSTADRATE
NII SAVASFRFSRCQPCAMENIEVAFVFGDNGNHLDDORLSVLTNNIYOGAT
OHEAGRAYVVRGDIKTSNDEVIVNANNKGGPQGGIATIRKMPGADKQPVAT
GKALVHGS PNVIVAHGPNRSRLSNEBGDKLSEYMDIATIRNERTKXISPLST
GITVAGKDRVMQSLNHLFTAMDITDADITITCLDQWESRIKEAITREKSEVELTEDD
RPVIDELVRHPLSLAGRPYSTEGRVYVYLBGTQRHQAQVAEINAMVPKQEA
NEOICLVYLGESMNSIRSKCVESSEASPHITPCLCNVMTAEARVYRLMANKQEP
AVGSSFOLPKYRITGVOKIOSKSPVIFSGTVPAPAIHPKRFASVTEDPVQPPRIAP
RRPAPVPVPARIPSPCTSTNGSTTSIOSLGEYOSASAASGAEISUDVQSLTSPSA
TGPVYRSTSSLSLEOPTPTPVVABEIHASGSLMSITSTIGFETRPVSPSODSRPS
TPSASASHTFVDLITPDSVAILEDPSFSPPQFLSEKPIPAFTRVNNMSRSDITIK
PIPKRCQVXYTQPPGVARAISAEFDEFVRHSN"
47. 1642
/product="NS1"
1643. 4024
/product="NS2"
4025. 5680
/product="NS3"
5681. 5683
/note="stop codon between NS3 and NS4 genes"
5702. 7528
/note="unnamed protein product; Protein sequence is in
conflict with the conceptual translation"
/codon_start=1
/protein_id="CAA44847.1"
/db_xref="GI:1334981"
/db_xref="SPTREMBL:066581"
/translation="YIPSSSEYGGHLOKSTROCKLQPILEERSVHEKFAVPLDLER
EKLQKLOLCASBGNRSYOSRKHENKALITVEKLLGIGSYISAEQPECKKVY
PAPMYSTASNSFSSAEVAVCNVLVLOENPTVASVNIITDEYADYLDMDVAGSCID
TATCPATLRSPFKHSYLRPEIRSAVSPIONTLQDVLAATATKRNCAVQTMRELPLV

DSAAFNVECFKKYCANDEYWDIFYTNPIRLTAENVTOYVTKLAKPKAAALFAKTHNLO
PLHEIPMDRFVMDLRDVKTGPTGKTEERKRVQVIOAADLALCGIARELVRRL
NAVLBNHITLFDMSADPDALIAEHPOFGAVLETDIASDFKSDAIAANSALMIE
DLGVQDLNLIIEAARGNITSYHLPTCTREFGMSMGMLTPIFIVNNIMASRV
LREHITTSPCAATFGDNNIKGVTSDMLMERCACTMLMEVKTIIDAVGVKAPYFCGG
FIVVDQITGCAFRVADPLKRLFKGKPLDDDVDVRRRLAHDEAARMNIGITEEL
VKAVESRYEVNYSLLITALLTTLASSVNFHGRHGHITLYG"
5702. 7525
/product="NS4"
7526. 7528
/note="stop codon between unstructural and structural
genes"
7534
/note="start for 26S sub-RNA"
7592. 11317
/note="unnamed protein product"
/codon_start=1
/protein_id="CAA44845.1"
/db_xref="GI:59186"
/db_xref="SPTREMBL:066579"
/translation="MPPEYPTLVPPMAPINPMAYRDPNRRPRRRPPRLAOTEDL
RRSIANTTLKORANPPAGPARKKAPSLISLRKKRPPPPAKKQKRRKPRGROR
MCMLESDDKTFPIMLNQVNGYACVVGARVEKPLHVBGRIDNEQLAIKLRKAS1YDL
EYGDVPQCMSDITLOYSDXRPGFYNMHGAVQYENNRFTVPRGVGSGSGRBILDN
KGRVVAIVLGVNIGSRATLSVVTMNOKVYVQKTPGSEBWSLATVWCILANITFPC
DQPCMCPCYEKNPHETLMLEONYDSRAVDOLDAVKKMARTRRDDLTHFTQYKL
ARPYLADPCNCGSRCSPLAIEVRDHAAGVRIQTSANFGKTRDLDLAVNSPMN
GKTQKSIKINDLHRTSAPCSLVSHHEGYILAQCPGDTVYVGHDEGNRTCTVAHK
VEFRPVAREKRYRHPHGEVLEPCNRVYTHKADQGVYEMQPLVAOHSLSLISAKY
KITVPSGAQVYKCYKCDVREGTSSDYTTCTDVKOCRAVLDINKKVVNSGRLPRG
EGDFEKKELHPEVYPAKCIATLAPPLVEHKKRTILHLPDPHLLTTRSLGSDA
NPQWIERPTVTFVTVGSLGLETTWGNHPRKWAQESGSGNPHGMHEVYVYNR
VPLTITIGCTCAVLIIVWCVSTSWILCRIRNICTPTPKAPNAOVILLALLCIRP
TRADDTQVANTLVNMMNNQFMWOTLPLIALIYCMMLKLPFCGSAFLVLCALGA
AAETHVMPNKVGIYKALVERGPVAVHLOQLVNTRIILPSTNLEITCKYTKYKPA
SPVKKCGAQOCTSKSPHDIQOVFTVYFPMWGAFCFDEMTOMSEAVLSSEBEC
SIDHAKYKATGTOQAMVNIYGVSWRSADVYNSETPAKIDAKLIGPLSASVS
PFDKVVQVLEHVVNIOPEFGTCKASPGLOGRSTSDNLVANTMLKLOPAGIY
HTPTVQAPSGEHRKPRKGAIPNDVAPFGSILAEPLRANCAVPISTIDPDAPF
TRISERTVSDLEKKTITECTYASPGIATVAYSSAGCNPHISPGVAVIKENDYT
LAEBSGTFHFSTHINHPARFLQVCTSAVTCGDKCPKHIVDPYPMQHTSFSAIS
ATAWSKVLVGVGTSAPRIVGLIATAVALVLPFHRH"
7592. 8371
/product="c protein"
8372. 8560
/product="B3 protein"
8561. 9820
/product="B2 protein"
9821. 9991
/product="6K protein"
9992. 11314
/product="E1 protein"
BASE COUNT 3322 a 2882 c 2805 g 2666 t
ORIGIN
Query Match 54.9%; Score 2745.2; DB 14; Length 11675;
Best Local Similarity 72.3%; Pred. No. 0; Mismatches 1368; Indels 6; Gaps 2;
Matches 3592; Conservative 0;
OY 14 ATGCATCCATATGGAAGAATTCACGTTGACTTAGATGCTGACAGCCGATGTCAAGT 73
DB 36 ACCATTCMAAATGAGAAAGTTCATGTGACTTAGACGACGACAGCCCATTCGTCAGT 95
OY 74 CATTACAGCGACGCTTTCCACAAATTTAGATGAGAGCAGGACGCTCACTGACAATGACC 133
DB 96 CACTGCAAGAGTGGCTTTCCACATTTGAGATGAGAGCAACGACAGTCACTGACAATGACC 155
OY 134 ATGCATGCGACAGCGCTTTTGGCATGAGCAACAAGCTCATGAGAGGAAGTGCAGC 193
DB 156 ATGCTAATGCTAGGCGCTTTTCGACCTAGCTACTTAACTCATGAGAGGAAGTGCAGT 215
OY 194 GGAACCAAGTTATCTTGACATTGGAAGTGGCCGCTCAGACATGACATTCCAATGACC 253

Db 216 CAGACAGATGATCTGGATATTTGGAGCGCGCTGTAAAGGCATACGCATCCAAACATA 275
Qy 254 GCTATCATTTGATCTGCCCTTATGATTAAGCGCTGAAGACCCGGACAGACTACAAAGGTATG 313
Db 276 AGTACCACTGATTTTGGCCCAATGAAAGCGCAGAAAGACCTGACAGACTTCAACGCTAAG 335
Qy 314 CAGAAAAGCTTAAAGAAAAGTACATTACCGACAAGAACATATAGCCTTAAAGCGCAGAC 373
Db 336 CAGACAAAGCTTAAAGAAAAGTATGTCACTGACAAATGATTTGCTCTTAAGGCCCGGAGAC 395
Qy 374 TGTGTGAAGTATGTCAACACCAAGACGCAAGACTCCATCTGTGTATGCAACAGACG 433
Db 396 TGTCAACAGTATGTCCAGCGCTGACCGCTGAGACACCTCGTATGATGACACTGACT 455
Qy 434 CCAGTGTAGGTACTTTTGAAGTATAGCATATACAAAGATGTGTACGCATCCATGAC 493
Db 456 CAACCTTGAGGTATCCAGCGCTCCGTGGCCGTATTCAGAGATGTATATGCACTGATGAC 515
Qy 494 CGACATCAATCTACACACGCGCTTAAAGAGTTAGACAATTTACTGATAGGCTTTG 553
Db 516 CGACTTCATTACTACAGCGCGCTGAAGGTGTACGAACATATCTATTGGATCGGGTTCC 575
Qy 554 ACACGACCCCTTTATATGTAACAAAAATGCGAGGTTCTTACCTTACTTAAACAGAACT 613
Db 576 ATACTACACCGTTCATGTATTAAGAACATGCGAGCGCTTACCTTACATACAAACAAATT 635
Qy 614 GGGCTGACGAGAGTATTTGGAACAGCTAACATTTGGCTCGGTATCTCAATCTTCAGG 673
Db 636 GGGCGGAGAAAGTGTGTGGAAGCAGAAATATATAGGCGTGGTATCAACTTGACG 695
Qy 674 AGACAGGCTTGAAGAACTCTCAATCCTTAGAAGAGGCTCCAACTTACTAATAAGA 733
Db 696 AAAAGAGTTTGGAAAAGTATCCATTTAGAGAAAGAAATTAACACCCATTAATAAG 755
Qy 734 TCATATTTCTCGTGTGTCAACATCTACACAGAAAGATATCACTGTATCGTAGCTGGC 793
Db 756 TAAATATTTTCTGTGGGGTCAACTATCTTACTGAAGAGAAATCTGTATCGCAGCTGGC 815
Qy 794 ATCTTCAAAGGTGTCCACTTGAAGAAAGTCTAATTCAAGGTATGATGTGGACCA 853
Db 816 ATCTACCTAATGTCTTTATCTTAAAGGTAAACCTAGCTTTACAGCGAGATGTAAACCA 875
Qy 854 TTTGACGCTGTGAAGGTAGCTCATCAAAAGATTAACATGACCCAGACATATACGTA 913
Db 876 TCGTCAAGTGTGGAAGGTATCGTTGTCAAGAAATTAACCTCATGCTCTGGATTTACGGGA 935
Qy 914 AAGTTGAGAACTTGGCGTCAACAATGATCGCAGAGGTTTCTTGAATTGCAAAAGTCAG 973
Db 936 AAGTGAATTAATCTTGCTTGCACATGACCGAGAGGATTTCTTAAAGTTGCAAGGTTACG 995
Qy 974 ATACGCTGCGCGGAGAGGTTTCTTTGCTGTGTGTACGTATGTATACCAACCACTTT 1033
Db 996 ACACGTTAAGAGGAGAGGCTCTTTTCCCGTGTGTACGTATGTGCAACGACGACACTGT 1055
Qy 1034 GCGATCAGATGACAGGATTTCTGGCACTGACGTATGTGATGATGACCAAAAACTAT 1093
Db 1056 GCGACAGATGACCGGGAATCTGGGCACTGACGTATGTGATGATGATGACCCCAAGAGCTGC 1115
Qy 1094 TGGTTGGGCTCAACCAAGATTGTGTCAATGTGAGAGCGCAAGAAATTAATAACAA 1153
Db 1116 TGGTTGGGCTCAACCAAGATTGTGTCAATGTGAGAACCAAGTAAACAAATTAACA 1175
Qy 1154 TGCAGAACTATCTATTAACAGTGTGTGCGCCAGAGCTTTTCCAGGTGGCGGTGAACATC 1213
Db 1176 TGCAGAAATTAATCTATTAACAGTGTGTGCGCCAGAGCTTTCTCGGGG---CCGGGAATAC 1232
Qy 1214 GTGCGCATTTGAGCAGAGAAAGAACTAGGGGTGCGGAGCGCACTCTTACTATGGGCT 1273
Db 1233 GCGCAGACTGTGAGAGCAAAAAAGGGCTTAGGGGTACGGGAAGCTTCCCTATGATGGGCT 1292
Qy 1274 GCTGCTGGGCTTTCAAGACCCAGAAATCAATCCATCTACAGAAAGCTGTGTACGCAA 1333
Db 1293 GCTGCTGGGCTTTCAAACTCAAGATCAATCATTTACAAAGAGACTGTGGAACTGAAA 1352

Qy 1334 CAATTAGAAAGTACCTGCCGCTTTGACTCATTTGTGATTTCCAGCGCTTACCAGCAGC 1393
Db 1353 CTATCAAGAAAGTGTCCCGCGCTATTTATTTCTTTGTCAATCCCAACCAACAGCTATG 1412
Qy 1394 GGTCTGATATGGGCTTCCGCGGTAGGCTCAAGCTGCTGTGAACCAATGTCAAAACCG 1453
Db 1413 GGCCTGATATAGGATTCGTGCGCAATTAAGATGTCTATTGACGCAAAAGAGGACCCG 1472
Qy 1454 CACCGCTATTAACAATGGCCGATGTGAGATCTGCGTGTCTTAACGCAAGAGCTGAG 1513
Db 1473 CTCCAATTAATTAATGAGGCGGACGTTGACACCTTAAAGGCTGCAAGATGAAGCTAAG 1532
Qy 1514 AAGTGTGCGAGCGAAGAGATCAAGAAAGCCCTGACACCTTGCTGCTCTGAATAAGAA 1573
Db 1533 CCGTGTGAGGCTGAGGAGCCGTATGTCAGCATTTACTCCATCTTCTCCGAGGTGCATA 1592
Qy 1574 AAGAGCCGTAGAGCGAAGTAGAAGCTCATTTATGCAAGAGCAGAGCAGTAGCGTGG 1633
Db 1593 AGAGAGCCGTAGAGGCTGATGATGACCTGATCATGACAGAGGCGAGCAGGCGTGG 1652
Qy 1634 AGACACACAGAGACATCATGAGGTGACAAAGTTACCAGGCGAAGAGAAATTTGGGCTT 1693
Db 1653 AGACACCTTAGACGACATCAAGGTCAAGATCCAGGGGAGAAATGATCGGCTGT 1712
Qy 1694 ACGCTATCTTTCAACCCAGCGGTATTTGAATAGTGAAGAACTGCGGTATCCACCCAT 1753
Db 1713 ACGCAGTCTTTTCAACCAAGCGGTCTTTAACAGCGAAGCTATGTTATTCACCCGT 1772
Qy 1754 TGGCGGAACAAGTATCTGTATATGCTCAACAAAGTAGGCGCAGAGATACAAAGTCAGC 1813
Db 1773 TAGCTGAGCAAGTCTGTGTATGCTCACAAGAGGCGCGCAGACGATACAAAGGTAGAG 1832
Qy 1814 CATACACGGAAGTATGTATCCAGAGAGGACGGGGTCCCTGTTCAGACTTCCAGG 1873
Db 1833 CATACACAGTATGATGTCTCCAGTGTATAGCTATACATCCCATTTCCAGG 1892
Qy 1874 CATGAGTGAAGAGGCTTACATCGTTTCAACGAGAGGAGTTGTGAACAGATACCTGC 1933
Db 1893 CTCTGATGAAGAGTGAACCAATAGTATTTAACAGAGCGAGTTCGTTAACCGTTACTTAC 1952
Qy 1934 ACCACATCGCAATCAACGAGAGCGGTAAACACTGACGAAGAGTACTATAGACTGTAA 1993
Db 1953 ACCACATTCGCGTTAACGAGAGGCAATTGAATACAGATGAAGAGTACTACAAAGGTGTGA 2012
Qy 1994 AGACTCAGACACAGACTTCAGAAATACGTCTTCCATATTGACGACGAAAGTGTGAAGC 2053
Db 2013 AAACCACTGAGACAGACTCTGAGTATGATTTGACATGACGCAAGAAAGTGTGTGAAGA 2072
Qy 2054 GAGAAAGCGCAGTCCCTTGCTTACCGGTATCTGTATGATCCACATTTGACAGAT 2113
Db 2073 AAGGGATGCGGACCAATGTGCTGTGGCGCAATGATTTGACCCGCCATTTCCACGAAT 2132
Qy 2114 TTTGCTACGAGAGTCTCAAGACAGACAGACAGCACTCAACAAAGTCCCAACATCGAG 2173
Db 2133 TTTGCTACGAGAGTTTAAAAACAGTCTGTGTGACCAACAAAGTCTTACATCGAG 2192
Qy 2174 TCTATGAGTGTCCAGGTTTCAAGTAAATCTGGAATCATCAAAAGGCTGTGACTAAGAAAG 2233
Db 2193 TCTACGAGTGTCCAGGTTTGTGGAAGTGTGATTAATCAAAAGGCTGTGTACCAAGGCTG 2252
Qy 2234 ATCTGTTGTGAGGCGAAGAGAAATGCGCAGAAATCATCAGGGATTAAGAGAGA 2293
Db 2253 ATCTGTTGTGAGGCGAAGAGAAATTTGATGAAATCATTAAGAGCTCAAAAGCA 2312
Qy 2294 TGAGACGTATGATGTTGCTGCTAGACCTGTCAATGATGCTTAAATGGGGTTAAGC 2353
Db 2313 TGGCGGAGATGATGATCGCGCCGCAAGTGAATGATGATGATGATGATGATGATGATGAT 2372
Qy 2354 ACCCGTTAATACCTGTATCATTTATGAGGATTTGCTGTGATGACAGGACGCTGTCTGG 2413
Db 2373 ACTCGTGCACACACTGTATCATAGAGAGGATTCGCTTGCATGACAGGACTCTGTCTAG 2432

OY	2414	CACGATGTCGATCGTCAAACTTAAGAAAGTGATATTGTGGGGGAGCCCAAAACAATCGC	2413
Db	2433	CACCTATGCGCATCGTCAAGCCAAAGAAAGTGATATTGTGGAGATCCGAAACATTCGC	2492
OY	2474	GCTTCTTAAACATGATGTGGCCGGAAGATCAATTTAACATGACATATGACCTGAAGGT	2533
Db	2493	GCTTCTTTAACATGATGTGTCTTAAAGTGCAATTTTAAACGAGATATGACAGAAAGTGT	2555
OY	2534	ACCATAAAGCATCTCTAGAGAGTGCACACAGACTGTAAACCGCATCGTCTCAAGCTCT	2593
Db	2553	ATCAAAAGATGTCTCTGGCGATGCACCTAAGACAGTGACATCCATGTGTTCCACCTGT	2612
OY	2594	TCTACGACAGAGGAATGAAAGACGGTTAACCTATGTGCTGATTAATATCATATGATACCA	2653
Db	2613	TCTAATGATPAACGATGTAGAAATGTGCAACCCATGCAATGATTAAGATCAATATAGTACCA	2672
OY	2654	CAGGGACCAAAAGCCGCGCAAAAGATGTGATTTCAACCTGTTTCAGAGATGGGNTA	2713
Db	2673	CCAGTACTACCAAACTTTAAAGATGACATATATTAACCTGTTCAGAGGGTGGGTTA	2732
OY	2714	AACAGCTACAGATTGACTTACAAAAAATCAAGAAATCATGATGATCGCGCTCATCGCAAGAC	2773
Db	2733	AGCAACTGCAAGTTGACTTACAAAGAACAGAGATCATATACCGCACAGCGGCTCACAGGGGT	2792
OY	2774	TTACGGCGAAAGCGCTTATATGCTGTACAGTACAAAGTACACGAGATCCACTTACTTCGC	2833
Db	2793	TTACTAGAAAAGGGGTATATACGACAGTGGCCCTCAAGGTCAATGAGAACCCGTTATACGCAC	2852
OY	2834	AGACTTCGAGACGCTGAACGTTACTTACACGACAGAAAAGCAAGCTTGTCTGGAAGA	2893
Db	2853	AGACACTGTGAGCATGTGAAATGTACTACTTACACGACAGAAAAGATATGATGAAAGA	2912
OY	2894	CGTATGCTGTGATCCCTGATTAAGACACTTACAGCTAAATATCCCGGGATTTTCAAGG	2953
Db	2913	CTTTGGCCGGTGAACCTTGATCAAGAGCTTGACAGCATCGTACCGGGTAAITTTCAACG	2972
OY	2954	CTTCAATTGAGAGACTGGCAGACGGCGCAACACGACGCCATTAATGGCACCGTCTTGATTAAGC	3013
Db	2973	CCACACTGGAAGAAATGGCAAGCTGACGTACGTATATGGCGAAAATTACTTGAGACAC	3032
OY	3014	CGCAGACAGCTGATGTGTTCAGAAATTAAGTGAAGTGTGCTGGCCGAAGCCTTAAAGC	3073
Db	3033	CAGCTAGCATGTACCTTTTCCAAAACAAAGTGAAGTGTGCTGGGCCCAACCGCTAGAAC	3092
OY	3074	CAGTCTTGGCCACCGCCCAACATTTGTGTCAGCAGACACACATGTGGAGACGTTTGACCCAT	3133
Db	3093	CTGTGTGGCCACCGCCCAATATTTACGCTGACCCCGTCCGACGTGGAGACTATTCAGAGGT	3152
OY	3134	TCAGCATGACAGACCGTATCTCACTGAAATGCGACTGAACCTTTTGTGACACAGGTTCT	3193
Db	3153	TCAAGGATGACAAAGCGTATTCGCTGAGATGGCCTTAACCTTTTGTGACACAGATTCT	3212
OY	3194	TTGAGATGACCTGCAAGTGGGTATTTTCCGCTCTACCGTGGCACTTACTTAAACAGGG	3253
Db	3213	TTGGGGTGCATTCACAGCGGGTGTCTCCGCGCCAACTGTTCGCTGCTTAAACACA	3272
OY	3254	ATCAGCACTGGGATTAATCTGCCAGGGGAAGAACATATATGGGCTTAATATGAGAGGTAGCAA	3313
Db	3273	ATGAACACTGGGATTAATAGCCAGGTCCAAACATATATATGGGTGTGCACTGGCCACGCTA	3332
OY	3314	AGGAGTGTCAACGGGATATCCGTGCATCAAAAAGCGGTTGACACAGGACGGGTAGCTG	3373
Db	3333	AAGAACTTGACAGTGGTATCTCTGTATTTGAAAGCGGTGGATTAAGGTAGAGTGGTGT	3392
OY	3374	ATATPAGGAATATATCCATCAAGACTTACTCTCCAAACAATTAATATGTGGTTCATTAATTC	3433
Db	3393	ACGTTTCGACAGACATATTAAGACTTAACCCCGCTAATAATGTGGTACCCCTGATATA	3452
OY	3434	GCCGGTTCGCCCATCTGTTGATCGTTGACCAAAAGGACAGGGTCAACATGATCAACAGCG	3493
Db	3453	GAAAGCTCCCACTCACTATTTGGTGTGCACACAGAGATACATCTGGAAACGGTATTAATCTCC	3512
OY	3494	GATTCCTATCTAAGATGAAGGCAATCTGTGTGGTGAATCGCGCATCTATACAGATTC	3553

Db	3513	AGCTAGTACCAGAAATGACAGAAAAAACCGTACTAGTAGTGAGTACACTTATGAAACATAC	3572
Oy	3554	CAGGGAAGAAAGTAGAGTCCATGGGTCCATTGGCCACTTAATACCATCAGGTGTATCTCG	3613
Db	3573	CAGGAAAGAGGTGAGACATTTAGGCCAACGCCACAATGACATTAATTAAGCGAACTGG	3632
Oy	3614	ATTGGGAATACCTAGCCATGTCCGTAAATATGACATTTATCTTTGTCAATGTTAGAACCC	3673
Db	3633	ACCTGGGCAATCTGCGCTTTAGGCAAAATATGACATCATTTTATTAACGTAGAGACTC	3692
Oy	3674	CGTACAGGAACCATTCATCCACACAGTCCGAGATCAGCTATTCACACACAGCATGCTTA	3733
Db	3693	CCTACCGCACACCATTTACCAACAGTCCGAGGACATGCGCATCAACACACAGCATGCTTA	3752
Oy	3724	CGTGAAGGCTGTCCACCACTGAAACAATGTCGCGAACTGTGTGGCTTTAGAGTATGGCG	3793
Db	3753	CTAGAAAAGCAGTGGACATTTGAAACAAAGCGGTACGTGCATGCAATTTGGGCTATGGGA	3812
Oy	3794	TTGCTGATCCGGCAACCGAAGATATCATCTGCGGTGGCAACGCTCATTTTAGTTTACC	3853
Db	3813	CTGGGGAACAGGCCACCGAAGACATTTATCTGCAAGTGGCTCGCTCATTTAGGTTCTAC	3872
Oy	3854	GTGTCTGTACCTTAAGAACATCGCCGAAATACTGAGGTTCTCTGTGTTCTTCCGCA	3913
Db	3873	GTGTGTGCCAACCGAAGTGTCCCTGGGAAAAACATGAGGTGCGCTGTGTTTTCGGCA	3932
Oy	3914	AGGACAAACGGCAACCAACATATACATACCGAGCACAACCTCGGTATGTGTTGACAAACATCT	3973
Db	3933	AGGACAAACGGCAACCATCTCCAGATCAAGTATGATGATGATGATGATTAACAAACATAT	3992
Oy	3974	ATCAAGGGTCAACCAAGTACAGGCGAGGAGAGCTCAGCCTACAGATGATCAGAGGTG	4033
Db	3993	ACCAAGGGTCAACTCAATCATATGAACTGGACAGACCTGGCTATAGAGTGGTGGCGCGG	4052
Oy	4034	ACATTAGCAAGGCGCTGACCAAGCTATCGTTAATGCTCTAATAGCAAGCTCAACGAG	4093
Db	4053	ACATTAACAAAGCAATGATAGGTTATTTGTTAACGGGCGCAACAAAGGCAACCGG	4112
Oy	4094	GTTCCGGAGTGTGGGGTGCATGTACCGAAATGGCGCGGCTTTGATATGACAGCA	4153
Db	4113	GTGGGGGTGTGTGGCGCCCTTTACAGAAATGGCTGGAGCTTTGATATGACAGCGCG	4172
Oy	4154	TAGCTGTCCGACCGGTAGACTTGTGGAAGCAGAACCGCTCATCATATCATGCTGTAGAC	4213
Db	4173	TAGCAACGTGTAAAGCGCACTGTTCAGACATTCCTCGAAGCTATCATGCGCTTGGCC	4232
Oy	4214	CCAAATTTTCTTAAGATCCCGGAACCGGAGGCGCACTTAAGCTCGACGTGCTTACATGA	4273
Db	4233	CTAATTTTCTCGGCTATCAGAAAAACGAGAGACAGAAATTTGTCGAAGTGTATACATGG	4292
Oy	4274	GCATAGCGTCCATGCTCAACGCTGAGGGGATTAACAAAATATACATACCGCTACTGTCAA	4333
Db	4293	ACATTGCGAATTAATCAACACGAGAGGTTTACTTAAGTCTCCATTCCTGTTTATCTTA	4352
Oy	4334	CCGGCATCTAATCTGTGTGGCAAAATGAGTGATGATCATCTGTCATCACTGTTCCTG	4393
Db	4353	CCGGCATTTACGACAGTGTGTAAAGACAGGGTTATGCAATGCTBAACCATTTATTACAG	4412
Oy	4394	CTTTGCACTTACGGAATGCCGATGTCACCATATTTGCTTGGATTAACATGGGAGACA	4453
Db	4413	CCATGTGATTAACGACGACGACATCATTAATTACTGTCTAGATTAAGAAATGGAGTCAA	4472
Oy	4454	GGATTAATCGAGCCATTCACCGCAAAAGAAAGCGTTCGAA--ATTCTGATGTATGACAGC	4510
Db	4473	GAATTAAGGAAGCTATACATCGAAGGAAGACGTTGAAGAGCTTACTGAGATTAACAGAC	4532
Oy	4511	CAGTAGACATTTGATGGTACGGGTCCACCAACAGCTCTTTGGACGCGACACAGTT	4570
Db	4533	CAGTTGACATTTGAACGTGTACGGGTACACCGCTTGACAGCTTGGCAGGTATGACCTGGTT	4592
Oy	4571	ACTCGTCAATGAGGCAAGTTGTATTTCAATCTGGAAGGTACAGATTCATCAACGCG	4630

GNLGRTRMYRMHGTIEIEIPLHVSHTSRPCHIVDGHGFLRLARCPAGDSITMEFKK
DSVTHSCVPEYEVKNPVRELYTHPEHGEAOQOVYAHQONGAJVEHMLPGSEV
DSIVHSJSGSVTVTPPAGTALVECEGKRISETITATFOVSCCTKECCBRALQ
NDKMYNSDKLPKAGATLKGKHLVFLAAGKCTVAPBPMITFGRSRSLKHPK
NPLYLTROLDPEHYTHELISEPARVHPITTEKMEVWENHNPKRMAOPBAPGN
HGLHEVITTHYHRYPMSTIIIGLSICAIIVSVASATWLNKSVSCULTEYRLPNA
RMPICLAVLCCARTARAEITWESLDHNNNQWEMIDLLPLAALIVTLKVCVC
VPLVLVAGAGAGAYEHATTMPSSQASISYNTIIVRAGYALPISITPTKILPYN
LEVYCHYKTMDSBPAIKCCSOECTPYRDEOCKVTGYVPEWGAAGYCECDENT
OVSAKYVMKSDCLADHAEYKAHTASVQALNITVGBHSIVTIVYVNGEPVANNV
KLTPGLSTAMPDPDKIYOYAGERIYNDPFEYAGGQAGADIOSRVSDDIYANT
NLVIOBPKAGAIHPYTOAPSGFQMKDMSLKFTRAPFCEGLITNTRARENCAVGS
IPLAFDIPDALFTRVSETPILSAECTNECVSSDFGIALIVKRSASKSKCAVHP
SGTATLKEAAVELTEOGSVTHIFSTANIHPEFRLOICTSYVTCCKDCDHPMDIVTHP
QYHAOTFPAASXTAMTWLSTLGGSAVITIIIGLIVATIVAMVYLTNQKNH"
BASE COUNT 3224 a 2825 c 2908 g 2484 t
ORIGIN
Query Match 42.7%; Score 2133.4; DB 14; Length 11441;
Best Local Similarity 64.7%; Pred. No. 0;
Matches 3229; Conservative 0; Mismatches 1736; Indels 24; Gaps 3;
QY 10 ACTAATGCATCCAAATGAGAAAGATTGACGTTGACTTATGATGCTGACAGCCCGATATGTC 69
Db 30 AATTACTTACCCAAATGAGAAAGATTGACGTTGACTTATGATGCTGACAGCCCGATATGTC 89
QY 70 AAGTCGTTACAGCGGACGTTTCCACAAATTTGAGATCGAAGCAAGGCGAGTCACTGACAAAT 129
Db 90 AGAGCTTTGCAAGGAGAGCTTCCCGAGTTTGAAGTGAAGAACCCAGAGCTCATGATATAT 149
QY 130 GACCATGCGCAATGCCAGAGCGTTTTCGATGTGCGCAACAAAGCTTATGAGAGCGCAATGC 189
Db 150 GACCATGCTAATGCGAGAGCGTTTTCGATGTGCGCAACAAAGCTTATGAGAGCGCAATGC 209
QY 190 GACCGGAGCCAGATTTATTTGACACTTGAAGTGGCGCCCTGACAGATGCACATTTCCAT 249
Db 210 GACCATGCGCAAGCATCTTGAACATTTGAAGTGGCGCCCTGACAGATGTATTTCTAAG 269
QY 250 CACCGCTATCATTTGATCTGCCCTATGATTAAGCGCTGAAGACCCGAGACAGATCAACAGG 309
Db 270 CATAGATTCATTTGATCTGTCCGATGATGATGTGCGGAAGATCCGAGACAGATTTGACAG 329
QY 310 TATGCAAGAAAGACTTAAAGAAAGT-----GACATTTACCGACAAAGACATGACCTCTAAG 363
Db 330 TATGCAACTAAGCTGAAGAAAGAAATTTGAAGAAATTAATGACAAAGGATTTGACAAAGAA 389
QY 364 GCGGCAAGCTGCTGGAAGTCAATGCAACACGACGAGAGATCTCATCTCTGTGTATG 423
Db 390 ATGAAGGAGCTGCGCGCGCTCATGAGCGACCTTGACCTGGAACCTGAACATATATGTGCTC 449
QY 424 CACAAGAGCGCAGCTGTAAGTACTTTGGAAGTGAAGATATCCAGATGTGTAAGCA 483
Db 450 CACAGCGAGAGTGTGTGCTGACAGAGGGGCAAGTGTGTTTACAGAGATGTAATCCCA 509
QY 484 GTTCATGACCGAGATCATCTATACCAAGCGGCTTAAAGAGATTAGACAAATTTATG 543
Db 510 GTTGACGAGCGGACAGTCTATATACCAAGCAACAAAGGAGTTAAGTGTGCTACTG 569
QY 544 ATAGGCTTTGACAGACCCCTTTTATGTAACAAAACATGCGAGGTTCTTACCCTTATAC 603
Db 570 ATAGGCTTTGACAGACCCCTTTTATGTAACAAAACATGCGAGGTTCTTACCCTTATAC 629
QY 604 AACAGCAACTGGGCTGACGAGAGATTTGGAAGCAGTAACTAGGCTGCGTAAGCA 663
Db 630 TCTACCACTGGGCGGAGAAACGCTGTTAACGGCTGTAACATAGGCTTATGACGCTCC 689
QY 664 GATCTTGAGAGAGCGGCTTGAACACTCTCAATCTTTAGAAAGAGGCTCCACCT 723
Db 690 GACGTCANTGAGCGGTACGTAGAGGAGATGTCATCTTTAGGAAGAAATATTTGAACCA 749
QY 724 ACTAATTAAGATCATTTCTCGGTGTGTTCAACATTTACACAGAGATGATGACTGTTA 783
Db 750 TCCAATATATGTCATATCTGTGTGCTGACCATCTACACAGAGAGGAGATTTACTG 809

QY 784 CGTAGCTGGCATCTTCCAAACGTGTTCATCTTGAAGAAAGTCTTAATCTTCACAGGTAGA 843
Db 810 AGGAGCTGGACACCTGCCCTCTGTATTTTACTTACCTTAGGTGCAAGCAAAATTAACATGTGG 869
QY 844 TGTGGAGCAATGTTCAGCTGTGAAGGTACGTCAATAAAAGATTAACGATCAGCCAGGA 903
Db 870 TGTGAGACTATAGTTAGTTGCGAGCGGGTACGTCGTTAAAGAAATAGCTATCAGTCCAGGC 929
QY 904 CTATAGCGTTAAGTTGAGAACTTGGCGCTCCCAATGCTATCCGAGGGTTTCTTGAGTTGC 963
Db 930 CTGATGGGAAGCCTTACAGCTATGCTGCCACGATGACCCCGAGGAGTTTGTGCTGC 989
QY 964 AAAGTCACAGATPACGCTGCGCGGAGAGGGTTCCTTTGTGTGTGTAAGATGTAACA 1023
Db 990 AAAGTCACAGATPACGCTGCGCGGAGAGGGTTCCTTTGTGTGTGTAAGATGTAACA 1049
QY 1024 GCCACACTTGGCATCAGATGACAGGATTTCTGCAACTGACGTTAGTGTGATGACGCA 1083
Db 1050 GCTACATGTGTGACCAAAATGACTGGCATCTGCGCAACAGATGTCAGTGGCGAGCG 1109
QY 1084 CAAAACATTTGTTGGGCTTAACCAAGAGATTGTCTCAATGTTAGAGACCAAAAGAAAT 1143
Db 1110 CAAAACATGCTGTGGGCTTCAACAGGCGCATGCTGCAACGCTGCGACCCAAAGAAAC 1169
QY 1144 ACTAACCAATGACAGAACTATCTATTAACCAAGTGGTCCGAGGGGTTTCCAGGTGGGCG 1203
Db 1170 ACCAATACATGAAGAAATTAATCTTTTCCGCTAGTGGCCAGGCAATTTGTAAGTGGCA 1229
QY 1204 CGTGAACATGCTGCCGACTTGAAGCAGAGAAAGAACTAGGGGTGCGGAGCGCACTCTT 1263
Db 1230 AAAGAAATTAAGAAAGATTCAGAAAGATGAGAGGCGCATAGGACTACGAGATAGACAGTTA 1289
QY 1264 ACTATGGCTGCTGCTGGGCTTTCAAGACCCAGAAATCAATCATCTTACAAAGAGCT 1323
Db 1290 GTTATGGGTGCTGCTGGGCTTTTGAAGGACAAAGATTAATCTATTTATTAAGCGCCA 1349
QY 1324 GGTACGCAAACTTAAGAAAGTACTGCGCTCTTTGACTATTTGATTCAGGCTT 1383
Db 1350 GATACCCAAACATCACTAAAGTAAACAGGATTTCCACTATTTGTGCTCCAGATTA 1409
QY 1384 ACCAGCAAGCGGCTCATATAGGCTTCCGCGTAGGCTCAAGCTGCTGTTGAACCACT 1443
Db 1410 GGCAGTAAACATGATGAGATGCGGCTGAGAACAAAGATTCAGGAAATGCTGAAGAGCGC 1469
QY 1444 GTCAAAACCGGACCGGCTATTTAATGAGCGGATGTGAGCATCTGCGTGTTCACAGCA 1503
Db 1470 AAGGAGCGGCACTCTCATTTACTGACGAGGACGTACAAAGAGCTAAGTGGCGAGCGAT 1529
QY 1504 GAACTGAAGAGTGGCTGACGCGGAGAGATTCAGAGAGCCCTGCAACCTTCTCCT 1563
Db 1530 GAGGCTAAGAGTGGCTGAGAGCGGAGAGTGGCGCTGACATCAACCTTTTGGCAGCT 1589
QY 1564 GAAATTAAGAAAGAGACCGTGAAGAGGAGAGTACCTCATTTATGCAAGAGGAGAGCA 1623
Db 1590 GATGTTAGAGAGCCCATCTGGAAGCGGATGTTGATGATGATGATGATGATGATGATGAT 1649
QY 1624 GGTAGCTGAGACACACAGAGACACATCAGGTTGACAAAGTTACCAAGCGGAGAGAG 1683
Db 1650 GGTAGCTGAGACACACCTGCTGCTGATTAAGGTTACAGCTATGCGGCGGAGAGAGAG 1709
QY 1684 ATTGGGCTTAAGCTATACCTTACCCAGGCGGCTATTTGAATAGTAAACCTGCGGT 1743
Db 1710 ATCGGCTTTTACGAGAGCTTTTCTCCAGAGCTTACTCAAGATGAGAACTATCTTGC 1769
QY 1744 ATCCACCATTGGCGGAGCAAGTACTGTATGACTCAAAAGGTAGGCGAGGAGATAC 1803
Db 1770 ATTCACTCTGCTGTAACAAAGTATAGTATTAACACACTGCGGAGAAAGCGGCTTAT 1829
QY 1804 AAAGTCAGCCATTCACCGGTAGGTCATTTGATACAGAGGAGCGGCGTCTGTTCAA 1863
Db 1830 GCCGTGAAACCTTACATGAGAAAGTGTGTGCTGACAGGAGACATGCAATACCGTCCAG 1889

QY	1864	GACITTCAGAGATGAGAGAGAGCGCTACAGATCGTTTTCACGAGAGGAACTTGGTAAAC	19233
Db	1890	GACITTTCAAGGCTCTGAGTGAAGAGCGCCACCATGTGTACCAAGCGAGATTTGGTAAAC	19499
QY	1924	AGATAACCTGCACCAATCGCAATCAACGAGAGGCGCTAAACCTGACGAGAGATCTAT	19833
Db	1950	AGGATATCTGCACCAATTTGGCCACACATGAGAGAGGCGTGAACACAGATGAAGAAATATTAC	20099
QY	1984	AAGA-CTGTAAGACTCAGAGACACAGACTCAGAAATACGTTTCGATATTGACGACGAAAG	20433
Db	2010	AAAACGTCAAGCCACGACGACACGACGCGCAAAATCCTGTATTGACATCGACGAGAAACAG	20659
QY	2044	TGTTTAAAGCCGAAAGAGCGCAGGTCCTTGTGTGCTTAACCGGTGATCTGTGTAGATTCACCA	21033
Db	2070	TGCGTCAAGAAAGAACTAATGTCACAGGGCGTAAAGGCGTAAAGCGAGCTGTGATCTTCC	21299
QY	2104	TTTTCACGAGTTTGCGTACGAGAGTCTCAAGACACGACGAGGACGCTCA-MAAGTCCCA	21633
Db	2130	TTTCATGAAATTCGCTTACGAGAGTCTGAAATAGCACCGGCGCTCTTACCAAGTACCA	21899
QY	2164	ACCATCGAGTCTATGAGTGCACAGTTCAGGTAAATCTGGAATCATCAAAAGCGCTGTG	22233
Db	2190	ACCATAGGGGTGTATGGCGTGCGCGGGAATCAGGCAAGCTGTGGCATATTAAAGCGCACTC	22499
QY	2224	ACTTAAGAAAGATCTGTGTTGTGTAGTGCAGAAAGGAAACCTGCGCAAAATCATCAGGAT	22833
Db	2250	ACCAAAAAAGATCTGTGTGTGTAGCGCCCAAGAAAGAAACCTGCGCGAAATATTAAGGAGC	23099
QY	2284	GTAAGAGAGATGAGAGCGTATGGAATGTGTCTGTAGGACTGTGCATTCAGTCTCTAAAT	23433
Db	2310	GTCAAGAAATGAAAGGCGCTGACGTCATGCTCAGAACCTGTGTGACTCAGTGTCTTGAAT	23659
QY	2344	GGGGTTAAGCACCCCGTTAACACTCTGTACATTGATGAGGCAATTTGCTGCGCATGACAGG	24033
Db	2370	GGATGCAAAACCCCGTGAAGCCCTGTATATTGACGAAGCTTTTGCTGTTCATGACAGGC	24299
QY	2404	ACGCTGCTGCGACATGATTTGCCATGCTCAACCTTAAGAAAGTGTATTGTGTGGGGAGCCCA	24633
Db	2430	ACTCTCAGAGGCGCTCATGAGCCATCATCACTAAGACCTTAAGAAAGAGTGCCTGTGCGGGACCCA	24899
QY	2464	AAACAATGCGGCTCTTAAACATGATGAGCGTGAAGATCAATTTAACATGACATATAGC	25233
Db	2490	AAACAGTCCGGCTTTTTCATATGTATGTGCTTGAAGAGTGCATTTTAACAGAGAAATTGC	25499
QY	2524	ACTGAAGTGTACCAATTAAGAGATCTCTAGAGGTGCACACAGACTGTAAACCGCATCTCTC	25833
Db	2550	ACGCAAGTCTTCCACAAAAGCATCTCTGTGTGTGACCTAATATCTGTACATTTCCGTGCTC	26099
QY	2584	TCACGCTCTTCTACGACAGCGAATGAAGACGTTAACCCATGTGCTGATTAATCATC	26433
Db	2610	TCAACCTTGTTTTACGACAAAAGATGAAAGACGAAACCCGAAAGAGCTAAAGATTGAG	26659
QY	2644	ATAGATACCAAGGAGACCAAAAGCCGACAAAGATGATCTGATCTTAACCTGTTTCAGA	27033
Db	2670	ATTGACATCTACGAGATCTTAACCAAAAGCAGCGATCTCATTTCTACATTTGTTTCAGA	27299
QY	2704	GGATGGGTGAACACAGCTACAGATTGACTTACAAAATTCAGAAATCATGACTGCGGCTCA	27633
Db	2730	GGGTGGGTGAAGCAGTTGCCAAATAGATTATTAAGGCAACGAAATATATGACGCGCAGCTGCC	27899
QY	2764	TCGCAAGGACTTAACGCGGAAAGGCGTTTATGTGTCAAGTACAAAGTCAACGAGATTCAC	28233
Db	2790	TCTCAAGGGCTGACCCGTTAAAGGCGTGTATGCGCTTCCGTTACAAAGTGAATGAATAATCC	28499
QY	2824	CTTATCTGCGACATTTCTGAGCAGCTGAACGTTTACTTACACGACACGAAAAAGCATTT	28833
Db	2850	CTGTACGACCACTTCAGAACATGTGAACGCTCTTACGTGACCCCGACGAGACGATATC	29099
QY	2884	GCTGCGAAGACGCTAGCTGTGTATCCCTCGAATTAAGACACTTAACGCTAAATATCCCGG	29433
Db	2910	GTGTGAAAAACATAGCCGCGCATCATGATTAATAAAACATGACGCGCAAGTATCAGGG	29659
QY	2944	GATTTTCAACGGCTTCAATGAGCAGCTGCGACGCGGAAACGACGCCATTATGGCAGCGTT	30033

Db	2970	AATTTCACTGCCACGATAGAGAAATGGCAAGACAGATGAAGCCATCATGAGCCATC	3029
Qy	3004	CTTGATTAAGCCGACAGACAGCTGATGTGTTCAGAAATTAAGTGAACGTCTGCGGGCAG	3063
Db	3030	TTGAGAGAGACCCGACCCCTACGACGATTTTCCAAATAATAGGGAAACGTGTGTGGCCAA	3089
Qy	3064	GCTTTAAGACCAAGCTCTTGGCCACGGCCAACTTGTGCTGACAGACAGCACTGGACACG	3123
Db	3090	GCTTTGGTGGCCGGTACTGAAACCTGACAGGACATACATGACCACTGAAACAATGAACT	3149
Qy	3124	TTGCACCCATTCACAGCATGACAGACGCGATCACTCACTTAATAGCAGCACTTCTTTG	3183
Db	3150	GTGAATTAATTTCCAAACGGACAAAGCTCACTGACAGATAGTATGTAATGAACCACTATG	3209
Qy	3184	ACCAAGTTCCTTTGGAGTAGACCTTGACAGTGGATTATTTCCGCTCTACCGTGCACTT	3243
Db	3210	GTGAGGTTCTTTGGACCTGACCTGGAACCGGGTCTATTTCTGCAACCACTGTTCCGTTA	3269
Qy	3244	ACTTACAGGGATCAACACTGGGATTAACCTCCGACGGAAACATGTATGGGCTTAATAGA	3303
Db	3270	TCCATTGGAGATATCACTGGGATTAATCCCGCTGCGCTTAACATGTACGGGCTGAATAAA	3329
Qy	3304	GAGGTACCAAGAGTGTGCAACGGCGAATTCGGTACATCAAAAAGGGGTGACACAGGC	3363
Db	3330	GAAGTGGTCGCTGACGCTCTCCCGAGGTACCAAGCACTGCTCGAGACAGTGTGCTACGGA	3389
Qy	3364	AGGGTACCTGATATTAAGGAATTAATACCATCAAGACATCTCCCAAAATTAATGTGGTT	3423
Db	3390	AGAGTCTATGACATGAACACTGGGACACTGGGCAATTAATATCGCGCATTAATCTAGTA	3449
Qy	3424	CCATTAAATGCGCGGTGCCCCACTCGTGTACTGTAACCAAAAGACAGGGTAACT	3483
Db	3450	CCAGTGAACAGAAAGACTGCGCTCATGCTTTAGTCTCCACCGTAAATGAACCCACAGAGT	3509
Qy	3484	GATCAACAGCGGATTCCTATCTAATAATGAAGGGCAATCTGTGTGGATGATCGGGATCCT	3543
Db	3510	GACTTCTCTTATTGCTGACGAAACTGAAGGGCAAACTGTTGTGTGTGGGGAAGAG	3569
Qy	3544	ATCAGCAATTCAGAGGAAGAAAGTACGTCCATGCGGTCCATTGCCATTAATACATCAGG	3603
Db	3570	TTGTCCGTCACAGCAAAACGGTTGATTTGTGTGACACACAGCCTGAGGCTACTTCAGA	3629
Qy	3604	TGTGATCTCGATTTGGGAAATCCTAAGCAATGTCGGTAATATGACATTAATCTTGTCAAT	3663
Db	3630	GCTCGGCTGGAATTTAGGATATCCAGGTGACGTGCCAAATTCGACATTTGATTAATTAAT	3689
Qy	3664	GTTAGACATCCCGTACAGGAACCATCTACCAACATGTCCAGGATCAACGCTATCCACAC	3723
Db	3690	GTCAGGACTCATTAATTAATATCAATCATTAATACAGATGTGAAGCAACGCGCATTAACCTT	3749
Qy	3724	AGCATGCTAAAGTGAAGGCTGTCCACCACTGAACACTGGCGGAACATGTGTGGCTATA	3783
Db	3750	AGTATGTTGACCAAGAAAGCTTGTCTGATTTGAATCCCGCGGAACCTGGGTACAGATA	3809
Qy	3784	GGGTATGGGCTTGGTGAATCGCGCAACCGAATAATCATCACTGGGTGGCACGCTCATTT	3843
Db	3810	GATTATGGTTACGCTGACAGGGCCACGAAAGCATTAATTTGGTGTATAGCCGGCAGTTTC	3869
Qy	3844	AGGTTTACCCGTGTCTGTACGCTTAAGAACATGCGGAAATATCTGAGGTTCTCTTGCTG	3903
Db	3870	AAAGTTCTCCCGGTATGCAAAACCAAAATCTCACTTAAGAGACAGAAAGTTCTGTTTGTG	3929
Qy	3904	TTCTTTGGCAAGGCAACGGGCAACACACATGATACAGGACAGACATCGGTTAGTGCTT	3963
Db	3930	TTCAATGGGTAAGATCCGAGGGCCGTAACGCAATCTTACAGCTTTATCATTAACCTTG	3989
Qy	3964	GACAAACATCTATCAAGGATCAACCAAGTACAGGACAGGAGAGCTCCAGCGTACAGAGTG	4023
Db	3990	ACCAACATCTATACAGGTTCCAGATCCAGAAACCGGATGCGCAACCTCATATCATGTG	4049
Qy	4024	ATCAGAGGTGACATTTAGCAAGAGCGCTGACCAAGCTATCGTTAATGCTGCTAATAGCAA	4083


```

Db      4050 GTGCGAGGGGACATTGCCACCGGCACCGAAGAGTATCATTAATGCTGCTAACAGCAAA 4109
Oy      4084 GGTCAACACGAGTTCCGAGTGTGCGGTGACTGTATCCGAAATGCGCGCTCTTTGAT 4143
Db      4110 GGACAACTGCGCGAGGGGTGTGCGAGCGCTGTATAGAAATTCGCCGAAAGCTTCGAT 4169
Oy      4144 AGACAGCAATAGCTGTGCGGACGCGTGAAGTGTGAAAGCAGAACCC--GCTCATCATTA 4200
Db      4170 TTACAGCCGATGAGAGTGGGAAAGCGGACGTGTCAAAAGGTGAGCTTAACATATCATT 4229
Oy      4201 CATGCTGTAGAGACCAATTTTCTTAAGATCCGGAAACCGGAGCGGACCTTAAGTCCGA 4260
Db      4230 CATGCCGAGGGCCCAACTTCAACAAAGTTTCGAAAGTTGAAAGGTGACAAACAGTTGGCA 4289
Oy      4261 GCTGCCCTCATGAGATAGCGTCCATTCGTCACAGCGTGAAGCGGATTAACAAAATATCAGTA 4320
Db      4290 GAGGCTTATGATCCATGCTTAAATTTGTCACAGTAAATTAACATTAACAGTCACTAGCATT 4349
Oy      4321 CCGCTACTGTCAACCGGACTTATCTGTGTGCAAAAGATCGAGTATGCAATCATTTGAT 4380
Db      4350 CCAGTGTGTCCACCGGATCTTTCCGGGAAACAAAGCCGACTAACCCATCATTTGAAC 4409
Oy      4381 CACCTGTCTACTGCTTTTGACACTAGCGGATGCGGATGTCACATATATTTCTTGATTA 4440
Db      4410 CATTTGCTGACAGCTTTAGACACCACTGATGAGATGTAGCCATATAGTCAGGAGCAAG 4469
Oy      4441 CAATGGAGACCGAGATATGAGGCAATTCACCGCAAGAAAGGCTGGAATTTCTGGA- 4499
Db      4470 AATATGGAAATGACTCTCAAGAGACAGTGTAGAGAAAGAGTGGAGAGATATGC 4529
Oy      4500 -----TGATGACAAAGCCAGTAGACATTGACTTGTGAGGCTCCACCCAAAC 4545
Db      4530 ATATCCGAGATTTCTTCCGTGACAGAACCGGATGAGAGCTGTGAGGTTATATCCGAAG 4589
Oy      4546 AGCTTTTGGAGGACAGACAGGTTACTTCTGCAATGAGGGCAAGTTGATCATACCTG 4605
Db      4590 AGCTTTTGGCTGAGAGAGGGCTACAGACAGATGATGCGAAGATTTCTCATATTTG 4649
Oy      4606 GAAGGTACAGATTCATCAGACCGGCCAAAGACATTGCCGAATTCATCCAAATGAGGCC 4665
Db      4650 GAAGGACCAATTTTACACGAGCGGCCCAAGGATTAAGCAAAATTAATGCTGAGGCC 4709
Oy      4666 AACCAATCTGAGGCTAATAGAGAGATTTGCTTGTATCTCTGGGAGAGATGTGTCAGC 4725
Db      4710 GTTGAAGAGGAGCCAAATGAGAGATATGCAATATCTTGTGTAAGAGATGAGCAGC 4769
Oy      4726 ATCCGCTCCAAATGCGCAGTAGAGAGTCAAGGCGTGTCTCACTTCAACACTTCCA 4785
Db      4770 ATTAGGTCAAAATGCGCGTCAAGAGTGGAAAGCTCCACACCACTAGACGCTGCCT 4829
Oy      4786 TGCCGTGTATTAATAGCTATGACGGGTGAGCGCGTATACAGGTGCGCTCTGCGAAGAA 4845
Db      4830 TGTCTGTATCTCAATGCTATGACTCCAGAAAGATCAACCGCTAAAGCCCTACAGTCCA 4889
Oy      4846 GAACAGTTCGCGGTATGCTATCTCTGTTGCGGAAAGTACAGAGTACAGAGCGGTGAG 4905
Db      4890 GAACAAATTAATCTGTGTCTCTCTTTCATTTGCGGAAATATGAAATCACTGTGTGAG 4949
Oy      4906 AAGCTACAGTACAGCAAAACAGTCTCTGTTTTCAGCGGTGTAACACCGGCTGTACACCC 4965
Db      4950 AAGATCCAGTGTCTCCAGCCTTACTGTTCTCACTAAGTCTGCGTACATTCATCCA 5009
Oy      4966 AGGAAGTAC 4974
Db      5010 CGGAAGTAC 5018

```

```

RESULT 7
AF004472 11420 bp RNA linear VRL 12-JUL-1999
LOCUS AF004472 Venezuelan equine encephalitis virus strain 66457, complete genome.
DEFINITION AF004472
ACCESSION AF004472.2 GI:5442461
VERSION

```

```

KEYWORDS
SOURCE
ORGANISM
Venezuelan equine encephalitis virus
Virus; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
Alphavirus; VEEV complex.
REFERENCE
1 (bases 8366 to 10039)
Powers,A.M., Oberste,M.S., Brault,A.C., Rico-Hesse,R.,
Schmura,S.M., Smith,J.F., Kang,W., Sweeney,W.P. and Weaver,S.C.
Repeated emergence of epidemic/epizootic Venezuelan equine
encephalitis from a single genotype of enzootic subtype ID virus
J. Virol. 71 (9), 6697-6705 (1997)
JOURNAL
MEDLINE
97404681
PUBMED
9261393
REFERENCE
2 (bases 1 to 11420)
Wang,E., Barrera,R., Boshell,J., Ferro,C., Freier,J.E.,
Navarro,J.C., Salas,R., Vasquez,C. and Weaver,S.C.
Genetic and phenotypic changes accompanying the emergence of
epizootic subtype IC Venezuelan equine encephalitis viruses from an
enzootic subtype ID progenitor
J. Virol. 73 (5), 4266-4271 (1999)
JOURNAL
MEDLINE
9214366
PUBMED
10196323
REFERENCE
3 (bases 8366 to 10039)
Powers,A.M., Oberste,M.S., Brault,A.C., Rico-Hesse,R.,
Schmura,S.M., Smith,J.F., Kang,W., Sweeney,W.P. and Weaver,S.C.
Direct Submission
Submitted (16-MAY-1997) Pathology, Univ. Texas Med. Branch,
Galveston, TX 77555-0609, USA
4 (bases 1 to 11420)
Wang,E. and Weaver,S.C.
Direct Submission
Submitted (12-JUL-1999) Pathology, Univ. Texas Med. Branch,
Galveston, TX 77555-0609, USA
REMARK
COMMENT
Sequence update by submitter
On Jul 12, 1999 this sequence version replaced gi:3849802.
FEATURES
location/Qualifiers
1..11420
/organism="Venezuelan equine encephalitis virus"
/mol_type="genomic RNA"
/db_xref="66457"
/db_xref="taxon:11036"
45..7502
/note="contains nonstructural proteins 1, 2, 3 and 4;
in-frame termination codon is read through at low
frequency"
/codon_start=1
/codon_except=(pos:5658..5660,aa:OTHER)
/product="nonstructural glycoprotein"
/protein_id="AAD43356.1"
/db_xref="gi:5442461"

```

CDS

```

/translation="MEKVVHVIDIEDSPFLRLAQSSFPPEVEBAKVYTNDAHNAFAS
HLAKLLETEDVPSDTILDIGSAPARRRYSHKHCHICPMKCAEDPDLVKYATLKK
NFKETIDKELDKKKELAAVMSDDLLETETCLHDDECRGQVAVQDVAVADGPT
SLHYDANKGVAVAYWIGFTTTPFMFKINAGAVPSSTWMADEVLTAAINIGSSDM
ERSRGMILKKTLLKSNVNVLFSGSTIYHEKDLBSMFLPSVFLHKGONVTRC
ETIYSCGVVKKRIALISGLYKXPSGVAATMHRGCTCCXDTDLNGBRSVPCTVY
PATCDQMTGILATDVSDADOKLVLGNQIVINVRQRTNTKATLPLVADAPA
RWAKEVYEDQEDERPLRLIRKMLERKESPLITIEDVOEACADAEVABEELRA
VLPRIGSNTLEIGLIRLIRKMLERKESPLITIEDVOEACADAEVABEELRA
ALPLADVEEPTLEADVLDLLOAGAGSVETPGLIVTSGVAGEDKIGSVASLPOA
VLKSEKSLCHIEPLAEVOVITHSRKGAVAEPPYKGVAVGAGHAIPODPOALES
TIVNBERFVNRVYLAHTATGCAANTDEYKIVPSRDEEYLIDRCKVKEIV
TGGLGTGEIVDPPEHFRAYESLRTTPAAPPVPTTGYGVGSGSGSIIKSAVTKDL
VVSARKENCAEIIIDVKKMKLDVNARTDVSVLNLNGCHPEVTLIDAEACVAGTFLR
ALIAIIRPKKAVLIGDPOKCGPFMMCKLAFHNEI CTQVPHKSISRCTGYSVSVS
TLFYDKMRTNTPKETKIEIDTGTSTXEDDLITLTCRGNVKLOIDIRKGNENTAA
ASOGITRGVAVAYKVNENPLVAPTSENVVLLRTEDRLVMKTLADGDPWIKTLTAK
YGNFTATIEEMOAEHAIIMRHIIERDPDTAPONKANVCAKALVLPDKTAGIDMT
EQMTVDVFETDKASHAEIYVNLQCYRFGIDLSGLPSATVPLSIRNNHWNDSPP
NMYGLANKSVQSSRSRIRPRRAVATGAVYDMNKTCLNDVPRNLVFNVRNLPLALV
LHNHEHPOSDFSFSKLGRTVAVVEKGVSTYMDWISDDEAFTFRALDGIPIG
AVPRYDVIIVNRTPYKYHNYOQCEBHAKIKLSMKACHLINFGCTGTSIGYADR
ASESITGAIARQFNSRVCKSPKSSLEETEVLVFLIGDRKARTNPNYLSLTINITYT

```


Db 1350 GATACCCAAACATCATCAAGTGAACAGCGATTTCCACTATTTGTGCTCCAGATA 1409
Qy 1384 ACCAGCCACGAGCTCGATATGAGCTTCCGCGTAGGCTCAAGCTGCTTGAACCACT 1443
Db 1410 GGCATTAACATTTGGAAGTTCGGGCTGAGAACAAAGATTCAGAAATGCTTAGAAGCGC 1469
Qy 1444 GTCAAAACCCGACCGGCTATTTCAAATGCGGATGTGAGCATCTGCGTGGCTTAGACAA 1503
Db 1470 AAGAGCCGTCACCTCTCATTTACTGCGGAGACGTACAAAGAGCTTAAGTGGCAGCGAT 1529
Qy 1504 GAAGGTGAAGAGTGGCTGCACCGGAGAGATCAAGAAAGCCCTGCCACTTGTCTCCT 1563
Db 1530 GAGGCTAAGAGAGTGCAGTGAAGCCGAGAGTGGCGCGTGCACCTACCTTTGGCAGCT 1589
Qy 1564 GAATATAGAAAAAGAGCCGTAGAGGACAGAACTGATTAATGAGAGGACAGAGCA 1623
Db 1590 GATGTGAGAGAGCCCATCTGTGAGAGCCAGATGTGACTTGAATCTACAGAGGCTGGGCGC 1649
Qy 1624 GGATGCGTGAAGACACACGAGAGACATCAGGGTGAACAAGTTACCCAGCGAGAGAG 1683
Db 1650 GGCCTAGTGAAGACACCTCGTGGCTTGAATAAAGTTACAGCTTAAGCGGCGAGAGCAAG 1709
Qy 1684 ATTGGCTCTTAAAGCTATCTTCAACCCGCGGCTTGAATATGATGAATGCGGCTGT 1743
Db 1710 ATCGGCTTTAGAGAGTGGCTTTCTCCAGGCTGTACTCAAGAGTGAAGAACTATCTTGC 1769
Qy 1744 ATCCACCATTTGGCGGAAACAACTAGCTGATGATGATCACTCAAAAGGTAAGGACAGATAC 1803
Db 1770 ATTCACTCTCTGCTGAACAACTCATAGTATACACACTCTGGCCGAAAGGCGGTAT 1829
Qy 1804 AAAGTGAAGCCATACACGAGTAAAGTTCATTTGACAGAAAGGAGCGGCTCCCTTTCA 1863
Db 1830 GCCGTGAAGCCCTACATGAGAAAGTATGCGCCAGAAAGACATGCAATACCCGTCAG 1889
Qy 1864 GACTTCCAGGCTTGAAGTGAAGCGCTACGATCGCTTTTCAACGAGAGGAGTTGCTAAC 1923
Db 1890 GACTTTCAGGCTCTAGTGAAGGTGCACCATTTGTGTCAACGAAACGAGAGTCTTAAAC 1949
Qy 1924 AGATACCTGCACCATCGCAATCAACGAGAGGCGCTTAAACATGAGAGAAAGTACAT 1983
Db 1950 AGGTATCTGCACATATTTGCCACACATGAGAGGCGCTGAACACAGATGAACATATATAC 2009
Qy 1984 AAGACTGTAAAGACTCAGAGACACAGACTCAGAAATGCTTTCATGATTTGACGACAGAAAG 2043
Db 2010 AAAATGTCAAGCCAGAGAGACAGACGCGCAATCTGTATGACATGACAGAGAAACAG 2069
Qy 2044 TGTGTAAAGCAGAAAGAGCGAGGCTCCCTTGTGCTTAAACGCGGTATGTGATTCACCA 2103
Db 2070 TGCGTCAAGAAAGAACTAGTCACTGGGCTTACAGCGAGAGCTGTGATCTCTCC 2129
Qy 2104 TTTCAGAGTTTGGGTAGAGAGTCTCAAGACAGACAGACGACCTCAAAATGCCA 2163
Db 2130 TTCATGATTTGCTTACGAGAGTCTGAGAACACAGCGCGCTCTTAAACCAATACCA 2189
Qy 2164 ACCATCGAGTCTATGAGAGTGCAGGTTCAAGGTAATCTGAATCATCAAAAGCGCTGTG 2223
Db 2190 ACCATAGAGGCTATAGCGGTGCGGCTTCAAGCAAGTCTGCAATATTAAGCCGCAATC 2249
Qy 2224 ACTAAGAAAGATCTGTTGTGTAGTGCAGAAAGAAACTGCGCAAAATCATCAGGAT 2283
Db 2250 ACCAAAAAAGATCTGTGTGTAGCGCCCAAGAAAGAAACTGCGCAAAATTAAGGGAC 2309
Qy 2284 GTAAGAGAGATGAGACGTATGAGATTTGCTGTAGAGATCTGTGATTCAGTGTCTTAAT 2343
Db 2310 GTCAAGAAATGAAGAGCGCTGAGCTCAATGCAAACTGTGACTCAAGTCTTTGAT 2369
Qy 2344 GGGGTTAAGCAACCCGTTAACTCTGTACATTTGAGAGCAATTTGCTGCATGACAGG 2403
Db 2370 GGATGCAAAACACCCGCTAGAGCCCTGTATATTTGAGAGCTTTTGTGTCTATGACAGG 2429
Qy 2404 ACGCTGTGCGCATGATTCGCTCAAACTTAAGAAAGTGTATTTGTGCGGAGACCA 2463
Db 2430 ACTTCAGAGCGCTCATAGCCATCAATGAACCTAAAGAGCGAGTGTCTGCGGAGATCA 2489

Qy 2464 AAACATGCGGCTTCTTAAACATGATGCTGAAAGTACATTTTAACATGACATATGC 2523
Db 2490 AAACATGCGGCTTCTTAAATATGATGCTGAAAGTACATTTTAACATGACATATGC 2549
Qy 2524 ACTGAAGTGAACATTAAGACATCTCTAGAGGTGCACAGACTGAACCGCATCTGC 2583
Db 2550 ACGAGGCTCTTCAAAAGACATCTCTGCTGCTCACTAAATCTGTGACTTCGGTCTGC 2609
Qy 2584 TCCAGGCTCTTCAAGACAGGAAATGAAGACGTTAACCCATGTGCTGATTAATCATC 2643
Db 2610 TCAACCTGTTTAAAGAAAGAAATGAGAACGAGAACCCGAAAGACCTAAGATTGAG 2669
Qy 2644 ATAGATACCAAGGACCAAAAGCCGACAAAGATGATCTGATCTTAACCTGTTTCA 2703
Db 2670 ATTGACATCTGCGAGATCTTAACCTGAGAGAGATCTCATCTTCTGTTTTCAGA 2729
Qy 2704 GGATGGGTGAACAGCTACAGATTGACTCAAAATTCAGAAATCATGACTGCGCTGCA 2763
Db 2730 GGGTGGTGAAGCAGATTCAATATGATTAACAAAGCAAGAAATATGACCGCAGCTGCC 2789
Qy 2764 TCGAAGACCTTACCGGAAAGCGTTATGTCTGAGTACAAAGTCAACAGAAATCCA 2823
Db 2790 TCTCAAGGCTGACCCGTAAGGCGGTATGCGCTTCCGTTACAGATGATGAATCCC 2849
Qy 2824 CTCTACGCAAGCTTCTGAGACGTGAACGTTACTTACACGCAAGAAAGACAT 2883
Db 2850 CTGTACGACCCACTTCAAAACATGTGAACGTCTACTGACCCGACGAGAGACGATAC 2909
Qy 2884 GTCTGAAGAGCTAGTGTGATCTCTGATTAAGACATTTACAGTAAATATCCGGG 2943
Db 2910 GTGTGAAGAAACATAGCGGGGATCATGATTAAGAAACATGACGGCCAGATCTCTGG 2969
Qy 2944 GATTTACGCGCTTACATGAGACATGCGAGCGCAACAGACCATTAATGSCAGCCTT 3003
Db 2970 AATTTCATGCGCAGCATGAGAAATGCGAGAGCATGAGGCCATCATGAGGACATC 3029
Qy 3004 CTGTGAAGCGGACAGACGTGATGTTTCCAGAAATTAAGTGAAGCTGCTGCGGCGAAG 3063
Db 3030 TTGAGAGACCGGACCTTACGACCATTTCAAGAAATTAAGGAGAGAGTGTGTTGGCCAG 3089
Qy 3064 GCTTTAGACGAGCTTGTGGCCACCGCCCAATGTGCTGACGAGACAGCAGTGGAGAG 3123
Db 3090 GCTTTGATGCGGATCTGAAGAACTGACAGGACATGACCACTGAACATGAACT 3149
Qy 3124 TTGACCATTTAAGCATGACAGAGCTTCACTGTAATGCACTGAACCTTTTTC 3183
Db 3150 GTGATTAATTTGAAACGAGAAAGCTCATCAGCAGAGATGATTAACCAACTATGC 3209
Qy 3184 ACCAGGCTCTTGAAGTGAACCTGACAGTGGGTTATTTCCGCTCTAACCTGCGACTT 3243
Db 3210 GTGAGTCTTTTGAACCTGACCTGACCTCGGCTCATTTTCTGACCCACCTGTTCCGTTA 3269
Qy 3244 ACTTACAGGATTCAGACCTGGGATTAATCGCCAGGAGAAACATGATGAGCTTAATGA 3303
Db 3270 TCCATTAAGAAATTAATCACTGGGATTAATCCCGTCCGTTAACTATAGCGGCTGAATAA 3329
Qy 3304 GAGGTAGCAAAAGAGTGTCAACGCGATATCTGTGATCACAAAAGCGTTGACACAGC 3363
Db 3330 GAAAGTGTCCGCTAGCTCTCCGCAAGTACCAAGACTGCTCGAGAGTGTCTACTGA 3389
Qy 3364 AGGAGTGAATTAAGAAATTAATCACTCAAGACTCTCTCAACATTAATGTGTT 3423
Db 3390 AGAGTATATGACATGAACACCGGACACTGTGCAATTAATGATCCGCAATTAATTAATA 3449
Qy 3424 CCATTAATTCGCGGTTGCCCACTGTTGATCTGTTGACCAAAAGACAGGCTTCAACT 3483
Db 3450 CCAAGTGAACAAAGACTGCTCAAGCTTTAATCTCCACATTAATGAACCAACAGAGT 3509
Qy 3484 GATCAACGCGATTTCTATCTTAAGTGAAGGCAATTTGTTGTGTATGCGGATCT 3543
Db 3510 GACTTTTCTTCAATTCGTGACGAAACTGAAGGCGAGAACTGTGTTGTGTGCGGAGAGAG 3569

CDS

```
/strain="66637"
/db_xref="taxon:11036"
45..7502
/note="contains nonstructural protease 1, 2, 3 and 4;
in-frame termination codon is read through at low
frequency"
/codon_start=1
/translation_except="(pos:5658..5660,aa:OTHER)
/product="nonstructural polyprotein"
/protein_id="AAD4335.1"
/db_xref="GI:5442460"
/translation="MEKVVHVIDEEDSPFLRLQNSFPQPEFAKQVNTDNDHANAFS
HLASKLIEVDPSDTLIDIGSAPARMYKSKHNCIPMCACEDPRLYKATLQK
NFKETIDELKMKMELAAMSDDPLETEMLHDESCREGQVAVQDYAVAGPT
SLYQANGKGVAVIIGDTPFPMKLAGAPSTYTMADETVLTANIGCSDDVM
ERSRGMGILRKLYLPSNNVLFSGSTIYHEKRLLSMPLPSFPHLRGMQNTC
ETVSCDGVVVKRIAISPLQYKPSGVAAHMBREPLCKDTLTLNRSVPCTTY
PATLCDMTGTLATVSDADOKLVLGNQRIVNGRTQRTNTMKNLTPVQAFA
RMAKEYKDEDEERPLGLRQLVNCGCWAPRRKITSYKPDQITIKNSDHSP
VLPRIGSNTLEIGLTRIRKMLEERKESPLIADVOEACADAKEAVEBELRA
ALPLADAVEEPTLEADVDMLOEAGASVETPRGLIYTSVAGEDKISYALSEA
VLKSEKSCIHPLAEQVITVHSGKGVAAEPYHGVVEBGAHPIPDQNALSESA
TIYNEREPNRYLHIHATHGALNTDEYKIVPSHDEEYLIDIRKQCKEYELV
TGLGTGLVDPPEHEFAESLRTAPAPVQVPTIGVPSGSGSGLIKSAVTKDL
VSAKKNCAEIRDVKKMGLDVNAVTVDSVLNGCHPEVETIYIDRAFAAGTLR
ALAIIRPKKAVLCGDPKQCGFNNMKCKVHNEICTQVHKSISRCTSIVSVS
TLFDKMRNTNPKETKIEIDTSGTKEODDLITCRGWKOLIDYKNEIMTAA
ASOGLTRGVAVRYKVNENPLIYAPTSBVAVLIRTRDITVMTLADGPKITLAK
YQNTETATIEEMQHEAMHRIILPRPTPAFOKANVCAKALVPLKTAGIDMTT
EOMTVDFETDKHSAEIVNQLCVRFGLDLSGLSAPLVLISINNNDMSPP
NMYGKNEVVRQLSRRYPLPRAVATGVYDMNTGTLNCPRIILVNNRLPVALV
LHNEHPQSDPSFVSKLGRVTLVVGKLSVPGKTVMLSDQEPATPRALDGLPG
AAPEKDIIVNRYRKYKHQOCEHAIKLSMLTKKCLHLPDGTCSVIGYAVR
ASESIGAIROPKSRVCKDKSLSEIETVFLPIGDRKATNHPYLSTIYNTV
GSRLEHGCASVHYRGDIATATGCVIMANSGQGGVCCALYKKPESFDLP
LEVGKRLVKAQAKIHAQVGENKVSVEGDKQLAAVESIAIIVDNNYKSAIP
LSTGIFSGKDRQLQSLNHLTALDITDVAVALCRKKEMLKEVAREVBEI
CISDSSVTEDEALIVHPSKSLAGRGYSTDGTSTYEGTFOHAARDIENA
MMPVATEANEOVCWYIIGESSISRSKQPVESASPTPLPCLCIAMPERQRL
KASRPEOITVSSPFLPKYRITGVOKIOCSOPIESPVPVPIHRIKIVTTPREN
OSTEGTPEPAPVNDATRTTRPEPIIEEBEDSISLSDGPHOVLVEADHPPS
VSSSWSIPIHASEPVDLSLIDLEAGVSEASASTNSVIFASMEHVLPAVAP
TFRNPPHAPRTTRTPSLPASRSRTLSVTPANNVITREBELATPRASRVS
SRTSLVNPENPNRVTITREEEAFVAAOQCFDAGAYFSSDTEGELQOOSQVTL
SEVLERTELEISYAPRLDOKEBELRKLQNPAPRSYOSRANMKAITARI
LOGIHYLKAGKVCYRTLPVLYSSVNRABSPRYAECANMKENPVIASIT
CTIPRYAYLDMVDGASCCLDJFASCPAKLSYFKHSLYPTTRISAPSAIQNTLQ
VLAATKNCNVLTQRELPLVDSAAFNVECFKRYACNNEIYETREKNIIRLEEVN
YITKLKGPKAALPAKTHNLMLQDIPMDREFMDLKRDKVTPGTGTHEERPKYQVIO
AADPLATAYLCGIHRELVRNLAVLPIHILFDMSADPAIIAEHQPQDQVLETD
IASPDKSEDDAMALTMILEDLGVDAELTLIEAPGEISIHLPPTKPKFGAMK
SGMELTLEPVNTVINIVIASRVLRBLNTPCAAFIGDNIYKGVKSDMLADRCATL
NMEYKIIDAVYGEKAPYFCGFIILCDIVTGAICRVADLKLFLKGLKLAADDEDD
RRRLHESHTMNRVGLIPELCKAVESRYETVGSIIYVAMTITLASSVKSYSYLAGP
ITLVG"
7538..11305
/note="contains capsid, E3, E2, 6K and E1 proteins"
/codon_start=1
/product="structural polyprotein"
/protein_id="AAC71187.2"
/db_xref="GI:5442459"
/translation="MFPEQPMYPMQPMFPRNPFAARRRPPRTDPLFAMOVETLS
MANLTFKORADAPPEGPAPKPKREAPPOKGGGKGGKGGKKAATGPNPVOS
GNKKTKNKKPKRORMMKLESQTPPIIMLGKINGVAVGUGLFRMHYEGKIDND
VLAATKTKAKSKYDLEADVPONNRADPTKTHEKPYIYSHHGAVYENGRFLVPE
GVGAKGDSGRPLIDQGRVAIVLGGVNEGRSTALSVNMEKGVATYETBENCOWS
LVTMCLANVTFCAPPCICYDRKAPTALMLSVNNDGSGDELELAIVCPGKGR
STEBLFEKYLTPYMARICICAVGSCSPALIAEVKSDGVDGRLDSTOYGDSS
GNLKGRTWRVDMHGTEIEIPLHOVSLHRSCHIVDGHYGLTANRGAVENHILGSEY
NLVHSGCSVPYEKFNPGRELYTHPRPHGBOACOVYAHQANRGAVENHILGSEY
DSLSVLSGSSVYTPPAGTCTALVECEGCTKISTINTAKOFSCTKECCRAVLO
NDKRVNSDKLPKAGATLKGHLVPLLAGKCTVPLAPBPMITFGRSYSLKLHPK
NPYLTTRQLADEPHYTHLISEPARNFTYTEKMEFVWGNHBPQMAOETAPGNP
```

Query Match	Score	DB	Length	Matches	Conservative	Mismatches	Indels	Gap
Best Local Similarity	64.6%	Pred. No. 0						
Matches 3223								
BASE COUNT	3210 a	2817 c	2306 g	2487 t				
ORIGIN								
10	ACTAATGATCAATATGAAAGAAATTCAGTTGACTAGTCTGACAGCCCGTATGTC	69						
30	AATTACCTAACCAAAATGAGAAAGTTACGTTGACATGAGAAAGACGCCATTCCTC	89						
70	AAATGCTTACAGCGGAGCTTCCCAATTGAGATCGAAGCAAGGACAGTCACTGACAT	129						
90	AGAGCTTTCGCAACGAGCGCTCCCGAGTTGAGTAGAAGCAAGCAGTCACTATAT	149						
130	GACCATGCGCATGCGCAGAGCGTTTTCGATGCGCAACAAAGCTCATTTAGAGCGAAGTC	189						
150	GACCATGCTATGCGCAGAGCGTTTTCGATGCGCTTCAAACTGATGAAACGAGAGTG	209						
190	GACCGGAGCAAGTATCTTGGACATTTGGAAGTGGCCCGCTCAGACATGACATTTCAAT	249						
210	GACCATGCGCAGACGATCTTGAATTTGGAAGTGGCCCGCGCAGAAATGATTTCTAAG	269						
250	CACCGCTTATCTTGTATTTGCGCTATGATAGCGCTGAAGACCCGGAACAGACTACAAAGC	309						
270	CATAGATCTCTTCATCTGTCGATGAGTAGTGGGAGTCCGAGACGATTTGATCAAG	329						
310	TATGAGAAAGACTTAAGAAAGT-----GACATTACCGACAAAGACATGAGCCCTTAAG	363						
330	TATGCACTAAGCTGAGAAAGAAATTTCAAGAAATTACTGACAAAGAAATTGACAAAGAA	389						
364	GCGGAGACCTGCTGGAAGTATGTCAACACGACGACGAGACTCCATCTCTGTGTATG	423						
390	ATGAAGGAGCTCGCGCGCGTATAGAGGAGCCCTGACCTGGAACCTGAGACTATATGCGCTC	449						
424	CACACAGCGCCACGTTAGTACTTTGGAAGTGTAGCATTAACAAATGATGTGACCA	483						
450	CACACGACGAGTGTGTGCTACGAAAGGCGAGTGTGTTACAGAGATGTATACCA	509						
484	GTCATGACGACGACATCAATCAACACGAGCGCTTAAAGAGTTAGGACAAATTTACTGG	543						
510	GTTACGACGACGACAAAGCTTTATACCAAGCCAAAGGAGTTAGAGTGGCTTACTGG	569						
544	ATAGGCTTTGACAGACCCCTTTATGTACAAAAACATGAGAGTTCTTACCTTACTTAC	603						
570	ATAGGCTTTGACACACCCCTTTATGTAAAGAACTTGGCTGAGCATATTCATCATAC	629						
604	AAACGAACTGGCTGACAGAGAGTATGGAAGACGTAACATTTGGCTTGGTACTTCA	663						
630	TCTACCAACTGGGCGGACGAAACCGTGTAAACGCTCGTAAACATAGGCTTATGACAGTTCC	689						
664	GATCTTACGAGACGAGCTTGAAGAACTCTCATCTCTTGAAGAAAGAGCTCCAACT	723						
690	GACGTCATGAGCGGTACATGAGGAGTGTCTTCTTGAAGAAAGATTTTGAACCA	749						
724	ACTAATAGATCATATTCCTGGTGTGTTCAACATTTACACAGAAATAGATCACTGTTA	783						
750	TCCATATATGCTTATCTGTTGGCTGACCATTTACACGAGAAAGAGGACTTACTG	809						
784	CGTAGCTGGCATCTTCCAAAGTGTTCATTGAAAGAAAGTAACTTACAGAGTGA	843						
810	AGGAGCTGGACACCTGCGCTGTATTTCACTTACGTGGCAAGAAATTAACATGTCTGG	869						

OY	844	TTGTGGACCAATTGTCAGCTGTGGAAAGGTAAGCTATCAAAAGATTAACATACGCCACGA	903
Db	870	TGTAGACTAATAGTTAGTTGTGGACGGGTACGTCTGTAAATAATAGCTATACGTCCAGGC	929
OY	904	CTATACGGTAAAGTTGAGAACCTTGGCCGTCCAAATGCACTCCGAGGGTTTCTTAGTTGC	963
Db	930	CTGTATGGGAAGCCATACGGCTATAGTCTCCACAGATGACCCCGAGGAACTTTGTGTCTGC	989
OY	964	AAAGTCACAGATACGCTCGCGCGGAGAGGGTTTCTTTGTGTGTGACGTATGTACA	1023
Db	990	AAAGTGACAGACACTTGTAMCGGGGAGAGGGTCTCTTTTCCCGTGTGCACGTATGTGCCA	1049
OY	1024	GCCACACTTTCGGATCAGATGACAGAGGATTTCTGGCACTGACGTTAGTGTGATGACGCA	1083
Db	1050	GCTACATTGTGTGACCAATATACCTGGACTACTGGCAAGATGTACAGTGCAGACGCGC	1109
OY	1084	CAAAAACATTTGTTGGGCTCAACCCAAAGATTGTCGTCATGTGAGACCCAAAGAAAT	1143
Db	1110	CAAAAACGTGCTGTGGGCTCAACCAAGGCATAGTCGTCAACGGTCGACCCCAAGAAAC	1169
OY	1144	ACTAACACAAATGCAGAACTATCTATTAACAGTGTGTGCCCGAGCGTTTCCAGGTGGCG	1203
Db	1170	ACCAATACCATGAGAAATTAATCTTTTGGCCGTAAGTGGCCGACGACATTTGCTAAGTGGCA	1229
OY	1204	CGTGAAACTCGTCCCGCACTTGGACGACGAGAAAGAACTAGGGGTGCGGAGACGCACTCTT	1263
Db	1230	AAAGAAATTTAAGAAATCAAGAAATATAGAGGCCACTAGACCTTACGAAATGACAGTTA	1289
OY	1264	ACTATGGGCTGTGTGGGCTTTTCAAGACCCAGAAATCACATCATTTACAAAGAGCT	1323
Db	1290	GTTATGGGTGTCTGTGGCTTTTAAAGGCACAGATTAATCTATTTATTAAGCGCCCA	1349
OY	1324	GGTACGCAAAATTTAAGAAATACCTGCGCTTTGACTCATTTGTGATTCACGCTT	1383
Db	1350	GATACCCAAACCATCATCAAAAGTGAACAGCATTTTCCATCTATTTGTGTCTGCCAGATA	1409
OY	1384	ACCAGCCACGGGCTGATATGGGCTTCGCCGTGAGCTCAAGCTGCTCTTGAACAACT	1443
Db	1410	GCGAGTTAACATTTGGAGATCGGGCTGTGAACAAGAAATCAGAAATTCCTGAAAGAGCGC	1469
OY	1444	GTCAAACCCGACCGGCTATTACAATGGCCGATGTGGAGCAATCTGCGGTTCACGCA	1503
Db	1470	AAAGAGCCGTCACTCTTAATTACTGCCAGAGCGTACAGAAAGCTTAAGTGGCAGCCGAT	1529
OY	1504	GAAGCTGAAGAAATGGCTGCACGCGAAGAGATCAGAGAAAGCCCTGCCACCTTGTCCT	1563
Db	1530	GAGGTTAAGAGGTGCTGTGAAGCGAGAGTGGCGCTGCACATTCACCTTTGGACCT	1589
OY	1564	GAATATGAAAAAGAGACCGTGAAGGCAAGATGACCTCATTTATGCAAGAGCGAGAGCA	1623
Db	1590	GATGTGTGAGAGCCCACTCTGGAAAGCCATGTGACTTGAATGCTTACAAGAAAGCTGGGGCC	1649
OY	1624	GGTAGCGTGAAGACCAACGAGGACCACTACGAGGAGCAAGTTTACCAGGGGGAAGAGAG	1683
Db	1650	GGCTAGTGAAGACACTCTGTGECTTGTAAATTAAGGTTACAGCTATGCCGGAGAGGACAG	1709
OY	1684	ATTGGGTCTTACGCTATATCTTCAACCCAGCGGATTTGAATAGTGAATAAACTGGCTGT	1743
Db	1710	ATCGGCTCTTACGAGTGTCTTCTCCACAGCTGTACTCAAGAGTGAATAAATCTATCTTGC	1769
OY	1744	ATCCACCCATTGGCGGAACAAGTACTGTGTAATGACTCAAAAGTTAGGCGCAGGAGATAC	1803
Db	1770	ATTCAACCTCTGCTGGAACAAGTCAATAGTATTAACACTCTGGCCGAAAAAGGGGCTTAT	1829
OY	1804	AAAGTCAGCCATTAACACGGTAAAGTCAATTGTATCCAGAAAGGAGCGGCGCTCCCTTTCAA	1863
Db	1830	GCCGTGGAACCTTACCATGAGAAATTAATGTGTGTCAGAGGACATGCAATACCCGCTCAG	1889
OY	1864	GACTTCCAGGCAATTAGTGAAGCCCTACGATGTTTTCAACGAGAGGAGTTCGTAAAC	1923
Db	1890	GACTTTCAAAGCTCTAGTGAAGAAAGTCCACATTTGTGTACACAGAACGAGAGTTTGTAAAC	1949
OY	1924	AGATATCTTGACACATCGCAATCAACGAGAGGCGTTAAACCTTGAAGAAAGTACTAT	1983

Db	1950	AGGTAATCTGCACCAATATTTGCCACATCGAGAGCCGTTGAACACAGATGAAGATATTTAC	2009
Qy	1984	AAGA CTGTAAGACTCAGAGACACAGACTCAGAAATACGTCCTTCGATATTGACGACGAAG	2043
Db	2010	AAAATTGTCAGAGCCACGAGAGACGAGCGGGCAATATTCGTATGATCATCGACAGGAAACAG	2069
Qy	2044	TGTGTAAAGCAGAAAGACGACAGTCCTTGTCCTAACCGGTGATCTGTGTAGATCCACCA	2103
Db	2070	TGCGTCAGAAAGAACTAATGTCATCTGGGCTAGGGCTTTCAGGCGAGCTGTGTGATCTTC	2129
Qy	2104	TTTCACGAGTTTGGGTACGAGAGTCGTCAAGACACGACACGAGCAGCACTCTCAAGTCCCA	2163
Db	2130	TTTCATGAAATTCGCTCTACGAGAGTCGTGAACACGACCGGCGCTCTTACCAAGTACCA	2189
Qy	2164	ACCATCGGAGTCATTGAGTGTCCAGGTTCCAGTAAATCTGGAAATCATCAAAAGCGCTGTG	2223
Db	2190	ACCATAGGGGCTGTATGGGCTGTCGGGTTACAGGCAAGTCGTGGCATTTAAAGCGCAGTC	2249
Qy	2224	ACTAAGAAAGATCTGCTGTGTGAGTCGCAAGAAAGAAAATTCGCGCAAAATCATCAGGGAT	2283
Db	2250	ACCAAAAAAGATCTGGGTGTGAGGCGCAAGAAAGAAAATTCGCGCAAGATTAATAGGAGC	2309
Qy	2284	GTAAGGAGATGAGACGTAATGGATGTTGCTGCTGAGACTGTCGATTCAGTCTCTTAAT	2343
Db	2310	GTCAGAAAAATGAAGGCTGAGAGTCGATCCAGAACTGTGACTCAGTCTCTTGAAAT	2369
Qy	2344	GGGGTTAAGCACCCCGTTAACACTCTGTGATCATTTGATGAGGCAATTTGCGCATGACAGG	2403
Db	2370	GGATGCAAAACCCCGTAGAGACCTGTATATTGACGAAGCTTTTGGCTGTCAAGCAGGC	2429
Qy	2404	ACGCTGCTGCAGTGAATTCCTCATCTGTCMAACTTAAGAAAGTGATATTGTGCGGGACCA	2463
Db	2430	ACTCTCAGAGGCTCATATGACCATCATTAAGACTTAAAGGAGGAGTCTCTGCGGGGATCCA	2489
Qy	2464	AAACAAAGCGGCTCTTTAACATGATAGCGCTGAAGATCAATTTAACATGACATATAGC	2523
Db	2490	AAACAGTGGGCTTTTTCATATATATATGTGCTGAAGAGCAATTTTAACACAGAAATTTGC	2549
Qy	2524	ACTGAAGTGTACCAATAAAGCATCTCTAGAGAGTGCAACACAGCTGTAAACCGCATGCTC	2583
Db	2550	ACGCAAGCTCTTCCACAAAAGCATCTCTGCTGTTGCACATAATCTGTGACCTTCCGCTGCTC	2609
Qy	2584	TCCAGGCTCTTCTACGACACAGCGAATGAGACGCTTAACCCATGTGCTGATTAATCATC	2643
Db	2610	TCAACCTTGTTTTACGACAAAGATGAGAACGACGAAACCCGAAAGAGACTTAAGATTGAG	2669
Qy	2644	ATAGATACCAAGGAGCCACAAAGCCGACAAABATGATCTGATTTCAACCTGTTTACA	2703
Db	2670	ATTGCACTACTGTGGAGTACTTAACCTTGAGCAGACGATCTCATTTCTGATTTTACA	2729
Qy	2704	GGATGGGTGAACACAGCTACAGATTGACTCAAAAATCACGAAATCATGACTGCGGCTGCA	2763
Db	2730	GGGTGGGTGAAGCAGTTGCCAATATGATTAACAAAGCAACGAAATATATGACGGCAGCTGCC	2789
Qy	2764	TCGCAAGGACTTACGCGGAAAAGCGGTTATGCTGTCAAGTACCAAGTCAACGAGATCCA	2823
Db	2790	TCTCAAGGGCTGACCCGTTAAAGCGGTGATCCGTTCCGTTCAAGGTGAATGAATAATCC	2849
Qy	2824	CTTACTGTGCAGACTTCTGAGCAGTGAAAGCTGTATCTTAACGCCACAGAAAAAGCATTT	2883
Db	2850	CTGTACGACCACTCTCAAGAACATGTGAACTCTCTACGACCCGCAACGAGAGACGTATC	2909
Qy	2884	GCTCGAAGACGCTAGCTGTGATCCCTCGAATTAAGACATTAACGCTAAATATCCGGG	2943
Db	2910	GTTGTGAAAAACATAGCCGCGGATTCATGATTAATAACATCTGACGGCCAAATATCTGGG	2969
Qy	2944	GATTTCAAGGCTTCAATTGACGAGCTGGCAGCGCAACGACGCGCATTTATGACGCGGTT	3003
Db	2970	AATTTCACTGCGCAGATAGAGAAATGGCAACAGAGCATGAGGCGCATCATGAGCAATC	3029
Qy	3004	CTTGATTAACCGCAGACAGCTGATGTGTTCCAGAAATAAAGTGAAACGTCTGCTGGCGAAG	3063

Dd 3030 TTGAGAGACCGGACCTTACCGACGATTCAGAAATAGCGCAACGTGTGTTGGCCAAAG 3089
Qy 3064 GCTTTAGAGCCAGTCTTGGCCACGCGCAACATTTGTGCTGACGAGACGAGCGAGGAGACG 3123
Dd 3090 GCTTTGTGTCGGTACTGAAAACCTGACGAGCATAGACATACCACTGAAATGAAACACTT 3149
Qy 3124 TTGCAACCATTCAGAGTACAGAGCGTATCTCACTGAAATGCACTGAATCTTCTTTTGC 3183
Dd 3150 GTGAGTTATTTGAAACGGAACAAAGCTCACTCAGCAGAGATAGTATGAAACCAACTATGC 3209
Qy 3164 ACCAGTTCTTTGAGTACGCTGACAGTGGGTTATTTTCCGCTCTTACCTGCGACTT 3243
Dd 3210 GTGAGTTCTTTGAGTACGCTGACAGTGGGTTATTTTCCGCTCTTACCTGCGACTT 3269
Qy 3244 ACTTACAGGATCAGACCTGGATATCTGCGAGGAAAGACATGATGGCTTATATGA 3303
Dd 3270 TCCATTAGGAATATACCTGGATATTTCCCGTCCCTTACATGTAACGCGCTGAAATGA 3329
Qy 3304 GAGGTAGCAAGAGTGTGACGCGGATATCCGTGCATACAAAAGCGGTGACACAGGC 3363
Dd 3330 GAAAGTCCGTCAGCTCTCCGCAAGGTACCCACGACTGCTCGAGCAGTTGCTACTGGA 3389
Qy 3364 AGGATAGCTGATATAGGAATATATCATCAAGACTACTTCCCAATTTATGTGTT 3423
Dd 3390 AGAGTCTATGACATGAAACCGGCACTGTGCAATATGATCCGCGCATTAATCTAGTA 3449
Qy 3424 CCATTAATATCGCGGTGCGCCCACTGTTGATCGTTGACCAAGAGGAGGATCAACT 3483
Dd 3450 CCAAGTACAGAAAGCTGCTATAGTTAGTCTTCCACATATAGAACCCACAGAGT 3509
Qy 3484 GATCAGACGGAATCTATATAGATGAAAGGCAATCTGTGTTGTTGCGGATCT 3543
Dd 3510 GACTTTTCTGCTGTCAGCAAACTGAAGGCGAGAACTGTTGTGTGCGGAGAAAG 3569
Qy 3544 ATCAGATTTCCAGGAAAGATGATGTCATGGTCCATTGCGCATATATTCATCAGG 3603
Dd 3570 TTGTCCGTCAGGAAACGCTGTGATCGTTGTGACACGACCTGAGGCTACTTTCAGA 3629
Qy 3604 TGTGATCTGATTTGGGAATACCTAGCCATGTGCGTAATATGATTAATCTTGTCCAT 3663
Dd 3630 GCTCGGCTGATTTAGTATCCAGGTGCGGCGCCAAATACGACATGTTATTTATAT 3689
Qy 3664 GTTAGACCCCGTACAGAAACCATCACTACAAAGTGGAGATCACTATCCACAC 3723
Dd 3690 GTGAGACTCCATATTAATATCATATTAATGAGCTGTGAAACACGCCATTTAAGCTT 3749
Qy 3724 AGCATGCTAATCTGTAAGGCTGTCCACACTGAAACCTGCGGAAATGTGTGCTATA 3783
Dd 3750 AGTATGTTGACCAAGAAAGCTGTGCTGATCTGAAATCCCGCGGAAACCTGCTCAGCAT 3809
Qy 3784 GGGTATGGGCTTGTGATCGCGCAACCGAATATCATCTGCGGTGGCAACGCTCATTT 3843
Dd 3810 GGTATGTTAGCTGACAGGGCCAGCAAGCAATTTATGTTGCTATAGCGGAGTTC 3869
Qy 3844 AGGTTTACCCGTGTCTGACAGCTTAAGAACTGCGGAAATAGTGGTTCTTCTG 3903
Dd 3870 AAGTTCTCCCGGATGCAAAACCAATCTCTGTAAGAGACAGAACTCTGTTTGA 3929
Qy 3904 TTCTTCCGCAAGGACAAAGGCAACACATGACACAGACAGTCCGTTGATGCTT 3963
Dd 3930 TTCTATTTGGGTACGATCGCAAGGCCCTGTACCAATCTTACATCTACCTTG 3989
Qy 3964 GACAACTCTATCAAGGGTCAACAGGTACAGAGGAGAGGCTCCAGCTACAGAGT 4023
Dd 3990 ACCAACTCTATACAGTTCCAGACTCCACGAAGCCGATGCGACCTCATATCATGTG 4049
Qy 4024 ATCAGAGTGAATATAGCAAGAGCCCTGACCAAGCTATCTGTAATGCTGCTATATG 4083
Dd 4050 GTGCGAGGGGACATTTGCCAGCGCCACGAGAGATGATATTAAGTCTGCTATACGCAA 4109
Qy 4084 GGTCAACCAAGTTCCGAGTGTGCGGTGACCTGTACCGAAATGCGCGGCTTTGAT 4143
Dd 4110 GGACAACTGTGGGAGGGGTGTGCGAGCGCTGTATTAAGAAATTCCTCGAAAGCTTGCAT 4169

Qy 4144 AGACAGCAATAGTGTGCGGAGCGCTAGACTTGTGAAGACGAAAC---GCTCATCTA 4200
Dd 4170 TTACAGCCGATCGAAGTGGAAAAAGCGGACTGTGTCMAAGGTGCGATATATCATT 4229
Qy 4201 CATGCTAGAGACCAATTTTCTAAGATGCCGGAACCGAGGCGACTTAAGCTCGCA 4260
Dd 4230 CATGCCGTAGGGCCCACTTCAACAAAGTTTCGAAAGTTGAAGGTGACAAACAGTTGGCA 4289
Qy 4261 GCTGCTACATGAGATAGCGTCCATGCTGCAACGCTGAGCGGATTAACAAAATATCAGTA 4320
Dd 4290 GAGGCTTATGAGTCAATGCTGTAATATGTCACGTAACATTAACATGCTACGAT 4349
Qy 4321 CCGTACTGTCAACCGGATCTATCTGTGCAAAAGTCGATGATGATCAATCATGCTAT 4380
Dd 4350 CCACGTGTGTCACCGGATCTTTTCGCGGAACAAAGTTCGATTAACCAATCATTTGAC 4409
Qy 4381 CACTGTCTACTGCTTTTGACACATAAGATGCGGATGTCACCATATATTTGCTGATTA 4440
Dd 4410 CATTTGCTGACAGCTTTTGAACACCACTGATGAGATGATGACATATATGCGAGGACAA 4469
Qy 4441 CAATGGAGACCAAGATATAGAGGCTTACCGCAAAAGAGTGGAAATTTGCTGAT 4500
Dd 4470 AAATGGAAATGACTCTCAAGAGACAGTGTGATGAGAGAAAGCATGAGAGATATGT 4529
Qy 4501 GATGCAAGCC-----AGTAGCATGACTTGTGAGGGTCCACCCAAAC 4545
Dd 4530 ATATCCGACGACTCTGCGTACAGAACCGATGAGAGCTGTGATAGGCTATCTCGAAG 4589
Qy 4546 AGCTTTTGGAGGACGACACAGTTACTCGTCAATAGAGGCAAGTTATTTATATACCTG 4605
Dd 4590 AGTTCTTTGGCTGGAAGAGAGGCTGACAGCAACAGCATGGAAGACTTTCTCATATTTG 4649
Qy 4606 GAAAGTACCAATTCATCAACGCGCAAGACATTTGCCAAATCATGCAATGTGCCC 4665
Dd 4650 GAAAGGACCAATTTCAACAGCGCGCCAAAGATATAGCAAAATTAATGCAATGTGCCC 4709
Qy 4666 AAACAAATGAGGCTAATAGACAGATTGCTTGTATCATCTCGGGGAGAGATATGCCAGC 4725
Dd 4710 GTTGCAAGGAGGCGCAATAGACAGATATGATATCTTGTGAAAGCATGAGCAGC 4769
Qy 4726 ATCCGCTTCAATGCGCCAGTAGAGAGTCAAGAGCGTGTGCTCCACTCACACTTCCA 4785
Dd 4770 ATTAGTCAAAATGCGCGTGAAGAGTGGAAAGCTCCACACACTATAGCAGCTGCT 4829
Qy 4786 TGCCTGTGTAATTAAGCTATACGCGTGAAGCGGTATACAGTTGCGTCTGCGAAGAA 4845
Dd 4830 TGCCTGTGATCATGCTATGATCTCAGAAAGATACAAAGCTTAAAGCTCACGTC 4889
Qy 4846 GAAAGTCCGCGATGATGCTCATCTCTGTTGCGAAGTACAGAGTACAGGCGGTGAG 4905
Dd 4890 GAACAATTAATCTGTGTGCTATCTCTTCAATGCGGAAAGTATGATATCATGTGTGAG 4949
Qy 4906 AAGTACAGTCCAGCAACAGTCTGTTTTCAGGCGTGTACCAACCGCTGTACACCCC 4965
Dd 4950 AAGATCCAGTCTCCAGCCTATACTGTCTCAGTAAGTGTGCTGCTATCATATCA 5009
Qy 4966 AGGAAGTAC 4974
Dd 5010 CGAAAGTAC 5018

RESULT 9
BEVCONGEN
LOCUS 11444 bp ss-RNA linear VRL 01-JUN-2001
DEFINITION Venezuelan equine encephalitis virus, complete genome.
ACCESSION L04653.1 L00931
VERSION L04653.1 GI:290609
KEYWORDS
SOURCE
ORGANISM
Venezuelan equine encephalitis virus
Venezuelan equine encephalitis virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
Alphavirus; VEEV complex.

REFERENCE 1 (bases 1 to 11444)
AUTHORS Kinney,R.M., Tsuchiya,K.R., Sneider,J.M. and Trent,D.W.
TITLE Genetic evidence that epizootic Venezuelan equine encephalitis
(VEE) viruses may have evolved from enzootic VEE subtype I-D virus
JOURNAL Virology 191 (2), 569-580 (1992)
MEDLINE 93079859
PUBMED 1448915
REFERENCE 2 (bases 1 to 11444)
AUTHORS Kinney,R.
TITLE Direct Submission
JOURNAL Submitted (11-JUL-1993) Molecular Virology, Centers for Disease
Control and Prevention, PO Box 2087, Fort Collins, CO 80522, USA
FEATURES
source
1. .11444
/organism="Venezuelan equine encephalitis virus"
/mol_type="genomic RNA"
/strain="P676"
/isolate="mosquitos/Venezuela/1963"
/db_xref="taxon:11036"
/tissue_1tb="puc18 and M13 of Kinney et al."
1. .11444
/function="genomic mRNA of VEE virus as determined from
strain P676"
1. .7526
/gene="NS"
1. .44
/gene="NS"
45. .7526
/note="possible incorporation of arginine, cysteine or
tryptophan at read through of UGA codon"
/codon_start=1
/trans_except=(pos:5682..5684,aa:Arg)
/product="putative nonstructural polyprotein precursor"
/protein_id="AAC19318.1"
/db_xref="GI:3201593"
/translation="MEKHVHDIEDSPFLRALORSFPQFEVKAQVTDNDHNAARS
HLASKLIEVDPSDTLIDIGAPARMYSKTHNCICPMKCAEDPRLTKATKCK
NCKEITDKEDLKKMELAAVMSDDLEETMCLHDSCEGQVAYQVYADGFT
SLHQAOKGVAVWIGFTDTPFMFKLAGAVPSYSTMADEVLTARNIGLCSDDM
ERSRGRMSILRKRYLKPSNNVLFVSGSTIYHEKDLRSWMLPSVFLRGKQNTCC
ETIVSCDGYVVKRLAISPLGKPSGVAATMREGPLCKVTDLNGSRVSPCTCY
PATLDDMTGLATDVASADOKLVLGNRIYVNGRTQNTMKYLLPYVAQARA
PMAKEYEDODERPLGLRDOLVMGCCWAFRRHKITSYKRPDTOTLIKNSDFHSF
VLPRIGSTLEIGRTIRKMLEHKEPSPLITAEIOEACADAEKVEAELELA
ALPPLADFEPTLEADVDLMLOAGASVETPGLIKVTSYAGEDKISYAVLSPOA
VLSEKSLCIHPLEAOVIIVTHSGRGRVAVEPYHGVVPEGAIIPODQVASEA
TIVNEREPVNRVYLHATHGALNTDEYKTVKPSBHDGYLIDRKQCVKELV
TGGLTGEIVLDPPEHFAVESLRTRPAPVQVPTIGYVPGSGSKSIKSAVTKDL
VSAKKENCAEIIIRDVKMKGLDVNARTVDSVLNGCKHPVETLYIDEAFCHAGTIR
ALAIATRPKAVILGDPKOCGFMMCKLVHFNHEICTQVPHKISIRCRKTSYTSVS
TLFYDKMRTTNPKETKIVIDTGSTKPKODLLITCRGVNKOLOIDYKGNEMTAA
ASQGLTRGVAVRKYVKNENPLVAPTEHVNALLTRREDRLVMTKLAGDPMIKILTAK
YGFNFTATIEBMOAENHAIIMRHLIERPDTDFONKANVCAKALVPLTKAGIDMT
EOMNTVDFETDKAHSAEIYLNQCFRFGDLDSGLFSAPTPLSIRNNHWDNSP
NMVTLKNEVNRQLSRYPOLPRAVATGTVDMNTGLRNDPRLNLPVNRRLPHALV
LHNEHQOSPFSSVSKLGRVYLVGEKLSVQKVDVMSDOPEATFARLDLIGTG
DVKRYDVFVNTPTPKYHHYOCEDEHAIKLSMTTKACILHNGGTCVSIIGYAADR
ASESIIGAIAROFKFSVCKPSKSHETEVLFVIGYDKARTINPKLSTLTNIT
ESRLHEAGCAPSVHVRGDIATATEGVIINANKSGOPGGVCGALYKPPESDLOP
IEVGRKALVGAAKHIIHAVGNPNKYSVEGDKOLAEYSAIKIYNDNNYSVALP
LSTGTFSGNKRDLTOSLNHLTLALDITDVAVALYCRDKKEMTLKEVAREAVEI
CISDSSVTEPDEALVPHKPSIAGRGYSTGQKPSYLEGKPHQAADIAEINA
MPVATEANEOVCNYIIAGESMSIRSKPVESBASITPSTLPCLCHAMTPEVOXU
KASRPEOITVCSFPPLPKYRITGVOKIQOSQPIIFSPRPVAYIHPKYLVEPTVEET
PESAPNOSTEGTEPEOPALVNDATRTMPEPIIIIEEBEESIIILSLGPHQVLOVE
ADHIGSVSSSSWSIIPHASDPVDLSIILDTLQASVTSPPGAVNATYREALEMR
ARVPAPRTVFRNPPHAPRTTRPPLAHASASRTISVTPGVNRTITREBELATP
SRASRSASRTISVSNPGVNRVITREFEFAVQAQORFPAGAYIIFSSDTGQCHLOO
KSVRQIVLSEVLERTELEISYAPRLDEKELIRKLOLNPTRANSRVEGSRVEMN
KAITARIILOGLHYLKAEGVECYRTLHPALPIISNRAFSPPKVAESCMLEK
NFPVASYCIIIPYDAYILDVADGASCCTIDFSPAKLRFPKSKSYLEPTIRASVS
AIQNTLQNVLAATKRNQNTVOMELPDLDSAAFVVECFKAYACNNEYWEFTKFNPIR

mat_peptide
/product="putative nonstructural protein nap1"
/function="minus strand RNA synthesis; methyltransferase;
guanylyltransferase"
1650..4031
/gene="NS"
/product="putative nonstructural protein nap2"
/function="RNA helicase; nonstructural proteinase;
necessary for subgenomic 26S mRNA synthesis"
/function="replication"
4032..5702
/gene="NS"
/product="putative nonstructural protein nap3"
/function="replication"
5703..7520
/gene="NS"
/product="putative nonstructural protein nap4"
/function="RNA polymerase"
/function="replication"
45..5684
/gene="NS"
/codon_start=1
/product="nonstructural polyprotein precursor"
/protein_id="AAC19320.1"
/db_xref="GI:3201594"
/translation="MEKHVHDIEDSPFLRALORSFPQFEVKAQVTDNDHNAARS
HLASKLIEVDPSDTLIDIGAPARMYSKTHNCICPMKCAEDPRLTKATKCK
NCKEITDKEDLKKMELAAVMSDDLEETMCLHDSCEGQVAYQVYADGFT
SLHQAOKGVAVWIGFTDTPFMFKLAGAVPSYSTMADEVLTARNIGLCSDDM
ERSRGRMSILRKRYLKPSNNVLFVSGSTIYHEKDLRSWMLPSVFLRGKQNTCC
ETIVSCDGYVVKRLAISPLGKPSGVAATMREGPLCKVTDLNGSRVSPCTCY
PATLDDMTGLATDVASADOKLVLGNRIYVNGRTQNTMKYLLPYVAQARA
PMAKEYEDODERPLGLRDOLVMGCCWAFRRHKITSYKRPDTOTLIKNSDFHSF
VLPRIGSTLEIGRTIRKMLEHKEPSPLITAEIOEACADAEKVEAELELA
ALPPLADFEPTLEADVDLMLOAGASVETPGLIKVTSYAGEDKISYAVLSPOA
VLSEKSLCIHPLEAOVIIVTHSGRGRVAVEPYHGVVPEGAIIPODQVASEA
TIVNEREPVNRVYLHATHGALNTDEYKTVKPSBHDGYLIDRKQCVKELV
TGGLTGEIVLDPPEHFAVESLRTRPAPVQVPTIGYVPGSGSKSIKSAVTKDL
VSAKKENCAEIIIRDVKMKGLDVNARTVDSVLNGCKHPVETLYIDEAFCHAGTIR
ALAIATRPKAVILGDPKOCGFMMCKLVHFNHEICTQVPHKISIRCRKTSYTSVS
TLFYDKMRTTNPKETKIVIDTGSTKPKODLLITCRGVNKOLOIDYKGNEMTAA
ASQGLTRGVAVRKYVKNENPLVAPTEHVNALLTRREDRLVMTKLAGDPMIKILTAK
YGFNFTATIEBMOAENHAIIMRHLIERPDTDFONKANVCAKALVPLTKAGIDMT
EOMNTVDFETDKAHSAEIYLNQCFRFGDLDSGLFSAPTPLSIRNNHWDNSP
NMVTLKNEVNRQLSRYPOLPRAVATGTVDMNTGLRNDPRLNLPVNRRLPHALV
LHNEHQOSPFSSVSKLGRVYLVGEKLSVQKVDVMSDOPEATFARLDLIGTG
DVKRYDVFVNTPTPKYHHYOCEDEHAIKLSMTTKACILHNGGTCVSIIGYAADR
ASESIIGAIAROFKFSVCKPSKSHETEVLFVIGYDKARTINPKLSTLTNIT
ESRLHEAGCAPSVHVRGDIATATEGVIINANKSGOPGGVCGALYKPPESDLOP
IEVGRKALVGAAKHIIHAVGNPNKYSVEGDKOLAEYSAIKIYNDNNYSVALP
LSTGTFSGNKRDLTOSLNHLTLALDITDVAVALYCRDKKEMTLKEVAREAVEI
CISDSSVTEPDEALVPHKPSIAGRGYSTGQKPSYLEGKPHQAADIAEINA
MPVATEANEOVCNYIIAGESMSIRSKPVESBASITPSTLPCLCHAMTPEVOXU
KASRPEOITVCSFPPLPKYRITGVOKIQOSQPIIFSPRPVAYIHPKYLVEPTVEET
PESAPNOSTEGTEPEOPALVNDATRTMPEPIIIIEEBEESIIILSLGPHQVLOVE
ADHIGSVSSSSWSIIPHASDPVDLSIILDTLQASVTSPPGAVNATYREALEMR
ARVPAPRTVFRNPPHAPRTTRPPLAHASASRTISVTPGVNRTITREBELATP
SRASRSASRTISVSNPGVNRVITREFEFAVQAQORFPAGAYIIFSSDTGQCHLOO
KSVRQIVLSEVLERTELEISYAPRLDEKELIRKLOLNPTRANSRVEGSRVEMN
KAITARIILOGLHYLKAEGVECYRTLHPALPIISNRAFSPPKVAESCMLEK
NFPVASYCIIIPYDAYILDVADGASCCTIDFSPAKLRFPKSKSYLEPTIRASVS
AIQNTLQNVLAATKRNQNTVOMELPDLDSAAFVVECFKAYACNNEYWEFTKFNPIR

misc_feature

/function="noncoding segment between nonstructural and

structural genes"

gene

/gene="26S mRNA"

mRNA

7532..11444


```

/gene="26S mRNA"
/product="26S mRNA region"
/note="The structural proteins of the virus are translated
as amino protein precursor molecule from an intracellular,
26S rRNA species that is identical to the 3'-one-third
portion of the viral genomic mRNA; putative"
/function="intracellular, subgenomic viral mRNA species"
7562..11329

```

```

/product="structural polypeptide precursor"
/protein_id="AAC19319.1"
/db_xref="GI:290611"
/translation="MFPPQPMYPMQPMPEYRNPFAAPRRPWFPRTPDEFLAMQVDELTRSMANLTFKORRARPAPGPPAKPKRKEAPKQKGGGKKKKKQKGGKKTAPPPAKQSGNKKKPKPKGROEMWKLLESDTKTFPLMEGLKNGVACVYGGLPDMYNGELINDVLAAKTKKASKYDLEVDVPPNNMRADTFKTHKPGQYYSMHGAINGYNGEPYAPKGVAGKSGSRPLINDQGRVYALUGVNGESRTLASVMNREKATYKTPPENCQMSLYTTMCLLANTPFCAEPPICIDYKRPHEATLMSVANDNPEDYLEDLAAKCPKRRKSTBELFKYKLTFRPYMARCIKAVGSCSPALIAVSDGSDVGLRTQSYQGLKPKSGNLKRRPYVMHGTEIEIPLHVOYLSHRECHVIDHGFLARCPAGSGIIMEFKGSHVSCSVPEYKVENIPEVGLYTHPEHGGAQCOVADQNGRVYVMHLPSEVDSLSLSSGSSVTPPYGTSALVCKCKGKGIETINKRKQSCCKKQCAVRLQNDKMYNSDXLPRKAAGLTGKGLHVPPLADGCTVPLAPMPITBFRSVSLAHKKNPTYLTQLADDEHYTHELISEPAVNFYTFEGKEFWMGNHPKRFMAQELPAGDPHGLPEHVEITHYRTPMSTILGSLCAALIVNSVAASVTLCKSRYSCLPYRLTPYARMPCLAVLACARTARAETWESLDHLMQNMQFMQLIPKSLIVTRLLKTVCCVVPFLVAVGAGAGAYEHATMPGQAGISYNTIVNRAGVAPLSIPITPKLIPYVNLLEYCHTKTGDSBPALCKCGSQCTPNRDEOCQKFTGYPFMGAGACPGDENTQSKRAYKMSDDCLADHGAAYKADHSAQVALINTVEHSLVITTVYNGGTFPYNFVCKLTKGISTAMTPDPDKRIVQYAGEIYVNDPEVAGQAGADQASVTSYSLYANTNLVYORPAGAHVPTYDASGFEQMKKDRAPSLKFTAPRCELIYNPITAEKCAVYSLPLAFEDIALFTFRISEITPYSAACTINECVSDSGGLATVYKASHPKCAVHPVSGTALKEAAVLETEQGSATHTFSTANIHPEFRVQISTVYTGCDGPKDHIYTHPQYHQATFPAASAKTAMWLTSLGGSANIIIGIYLAITIVAMVLYMOKHN"
7562. 8386

```

```

/product="core protein"
/function="encloses the genomic mRNA molecule"
/standard_name="capsid protein"
/note="putative"
8387..8563
/gene="26S mRNA"
/product="viral cleavage product"
/standard_name="E3 protein"
mat_peptide

```

Query Match	42.4%	Score 2122.2;	DB 14;	Length 11444;
Best Local Similarity	64.5%	Pred. No. 0;		
Matches 3225;	Conservative	0;	Mismatches 1748;	Indels 24; Gaps 3;

Oy	2	CCCTGCAAAAGTAATGCATGCAATATGGAAGAAATTACAGTTGACTTAATGCTGCACAC	61
Db	22	CCCAAAACCAATTACTCTACCCAAATATGAGAAAGTTACGTTGACATCGAGGAAGACAGCC	81
Oy	62	CGTATGTCAAAGTCGTTAGCAGCGGACGTTTCCAAATTTGATTCGAAGCAAGCAGCTCA	121
Db	82	CATTCCTCAGAGCTTTTACACAGGACGTTCCCGCAGTTGAGGTAAAGCAAGCAGCTCA	141
Oy	122	CTGACAAATGACCATCTCCAAATGCAGAGGTTTTGCGATGTGGCCAAAGCTCATTTGGA	181
Db	142	CTGATTAATGACCATCTTAATGCGAAGCGTTTTGCGATCTGGCTTCAAATCTGATCGAAA	201
Oy	182	GCGAAGTCGACCCGGGACCAAGTTATCTTGGACATTGGGAAGTGCGCCGTCAGACATGCA	241
Db	202	CGGAGGTGAGCCCATCCGACAGATCTCTTGACATTGGGAAGTGCGCCCGCCGACAAATGT	261
Oy	242	ATTCCAATCACCGCTCATGTTGATCTGCGCCCATATGATTAAGGCTGAAGCCCGGACAGAC	301
Db	262	ATTCTAACCATTAAGTATCATCTTGCACTCTGCGATGATGTGCGAAGATCCGACAGAT	321
Oy	302	TACAAAGGTATGACAGAAAGACTTAAGAAAAGT-----GACATTACCGACAAAGCATAG	355
Db	322	TGTACAAAGTATGCACTAAGCTGAAGAAAAATTGCAAGGAATTAATGACAAAGGAATTGG	381

QY	356	CCTCTAAGGCGGAGACCTGCTGGAAGTCATGTCAAACACAGACGAGAGCTCCATCTC	415
Db	382	ACAAAGAAATGAAGAGCTCCGCCGCTCATATAGCGACCTTACCTGGAACTGAGACTA	441
QY	416	TGTGATATCACAACAGACCCACAGTGTAGTACTTTTGAAGTGTAGCAGTATACCAAGT	475
Db	442	TGTGCTCCACGACGATGAGTATGTCGTATGAGAGGGGCAAGTCCGTGTTTACAGAGATG	501
QY	476	TGTACGCGTCCATGCACTGCACTCCGACATCAATCTACCAACGAGCGCTTAAAGAGTTAGACAA	535
Db	502	TATACGCGTTGACGCGACCGAACMACTCTCTATCACCAGACCAACAGGAGTTAGTGTG	561
QY	536	TTTACTGATPAGGCTTTGACACGACCCCTTTATATGACAAAAACAATGGCAGGTTCTTACC	595
Db	562	CCTACTGATATAGCTTTAGACACACCCCTTTTATATGTTTAAAGACTTGGCTGAGCATATC	621
QY	596	CTACTTACAAACGAACTGGGCTGACGAGAGATATTGGAMCACTGTAACATTTGGCCTCG	655
Db	622	CATCATACTTACCAACTGGGGCGACGAAACCGTGTAAACGGCTGTATCATATGAGCTTAT	681
QY	656	GTAATTCGATCTTACGAGAGACAGGCTTTGAAAACTCTCAATCCTTAGAAGAAAGGC	715
Db	682	GCAGCTCCGACGTCAATGAGCGGTCACGTAGAGGATGTCATTTCTTAAAGAAAGTATAT	741
QY	716	TCCAACTACTAATATAGATCATATCTCGGTTGTTCACAAATCTACACAGAAAGATAGAT	775
Db	742	TGAAACATCCAAATATATGTCATTTCTCTGTTGGCTGCACATCTACACAGAAAGAGG	801
QY	776	CAGTTTACGTAGCTGCGCATTTCCAAACGTGTTCCACTTGAAGAAAGTCTACTTCA	835
Db	802	ACTTACTAGAGAGCTGGGACCTGGCTGTATTTTCACTTACGTGGCAAGAAATATTAAC	861
QY	836	CAGGATAGATGGGACCAATTTGCAGCTGTGAAGGGATCGTCATCAAAAGATATACGATCA	895
Db	862	CATGTCGGTGTGACACTATAGTTAGTTGCGACGGGATCGTTAAAGAAATAGCTATCA	921
QY	896	GCCCGAGACTTACCGTAAAGTTGAGAACTTGGCGCTCCACAAATGCATCGCAGAGGTTTCT	955
Db	922	GTCCAGGCTGTATGGAGAACCTTCAGGCTATGCTGTAACATGACACCGCAGAGGATTC	981
QY	956	TGAGTTGAAAGTCAACAATACGCTGCGCGGAGAGAGGTTCTTTGCTGTGTGATCGT	101
Db	982	TGTGTGAAAGTGCACACACATTTGAACGGGAGAGGGTCTCTTTTCCCGTGTGACGT	104
QY	1016	ATGTACACGACACTTTTGGCATCGATGACAGGAGTTCTGGCAACTGACTGTAGTGTG	107
Db	1042	ATGTGCCAGCTACATTTGTGTACCAAAATGACTGGCATCTGGCAACAGATGTCAAGTCGG	110
QY	1076	ATGACGCACAAAATTAATTTGGTGGGCTCCAACCAAGGATTTGCTCAATGATGAGACGC	113
Db	1102	ACGACGCCCAAAAATCTGCTGGTTGGGCTCCAAACACGCAATATGCTCAACGGTGCACCC	116
QY	1136	AAAGAAATACTAACAACATGACGAACATATTAATTCACAGTGTGCGCCGAGGGTTTCCA	119
Db	1162	AAAGAAACCAATATTCATGAGAAATTAATCTTTTCCCGTATGTGGCCACAGGCAATTTGCTA	122
QY	1196	GGTGGCGCGTGAACATGTCGCCGACTTGGACGACGAGAAAGAACTAAGGGTGTCCGGAGC	125
Db	1222	GGTGGGCAAAAGAAATATTAAGGAAGTCAAGAAAGTAGAGGGCCACTAGAGCTACGAGATA	128
QY	1256	GCATCTTACTATGGGCTGTGCTGTGGGCTTTCAGACCCAGAAAATCAATCTCATCTTACA	131
Db	1282	GACAGTTATGTCATGGGCTGTGCTGGGCTTTTGAAGGACACAAGATATACATCTAATTTATA	134
QY	1316	AGAAAGCTGTATCGCAAACTAATTAAGAAAGTACTGCGCTTTTACTCATTTGTGATTC	137
Db	1342	AGCGCCAGATATCCCAACATCATCAAAAGTGAACAGGATTTTCCATCATTTCTGTGTGTC	140
QY	1376	CACGCTTTTACAGCCAGGGCTCGATATAGGCTTCCGCCGTATAGGCTCAAGCTGCTGCTTG	143
Db	1402	CCAGGATATGGAGATMACACTGTGAGATCGGGCTAGAACGAGATTCAGAAAAATGCTAG	146

Qy 1436 AACCACTGTCAAAACCGCACCGGCTATTACATATGCCGATGTAGAGATCTGCGTCT 1495
Db 1462 AAGAGCAACAAGAGACCTCTCTCATTTACTGGCGAGGACATACAAAGAGCTAAGTSCG 1521
Qy 1496 TACAGCAAGAAGCTGAGAAAGTGGCTGCAGCGGAAAGATCAGAAAGCCCTGCCACCT 1555
Db 1522 CAGCGGATGAGGCTAAGAGAAATGGCTGAAAGCCGAGGAGCTGCGGCTGCTTACACCTT 1581
Qy 1556 TGCTCCCGTAAATGAAAGAAAGACCCGTAGAGGAGAAAGTAGACTCTATTATGCAAGG 1615
Db 1582 TGGAGCTGATTTTGGAGAGCCCACTCTGGAAGCCGATGTGACTTGATTTTACAAAGG 1641
Qy 1616 CAGAGCAGGTAGCGGTGAGACACCAAGAGCAATCAGGGTGAACAATTACCAGCG 1675
Db 1642 CTGGGGCGGCTCAGTGGAGACCTGTGGCTTGATTAAGATTACCAAGCTATCCGCG 1701
Qy 1676 AAGAAAGATTGGGCTTTACGCTTACTTTTCAACCCAGCGGCTTATGAAATGTGAAAA 1735
Db 1702 AGGACAAGATGGGCTTTACCGCAGTGTCTTCCACAGGCTGTACTCAAGAGTGAAGAAC 1761
Qy 1736 TGGGCTGTATCCACCCATGGCGGAAACAAGTACTGTAAAGCTCAACAAGGTAGGCGAG 1795
Db 1762 TATTTGATTCACCTCTCTGCTGAACAAGTATGATTAACAACACTTGGCGAAAG 1821
Qy 1796 GAGAGTACAAAGTGCAGCCATACACGCTAAGTCTATTGTACCAAGAGGAGCGCGTCC 1855
Db 1822 GGGCTTATGCCGTGGAACCTTACCATGGAAGAAAGTGTGTGCGAGAGGAGCATGCATAC 1881
Qy 1856 CTGTTCAAGACTTCCAGGCAATTGAGTGAAGCGCTACGATGTTTCAAGAGAGGAGT 1915
Db 1882 CCGTCCAGGACTTTCAAGCTCTGAGTGAAGTGCACATGTTGACAAACAAGAGAGT 1941
Qy 1916 TCGTAAACAGATACCTGACACACATCGCAATCAACGAGGCGCTAAACCTGACGAAG 1975
Db 1942 TCGTAAACAGTACTGACACATATTTGCAACATGAGAGAGCGCTGAACCAAGTGAAG 2001
Qy 1976 AGTACTATAGACTGTAAAGACTCAGACACAGACTCAAGATAGCTTTCATATTTAG 2035
Db 2002 AATATTACAAAACGTGTCAAGGCCAGCGAGCAAGAGCGGAATACCTGTACACATTCGA 2061
Qy 2036 CACGAAAGTGTATTAAGCGAAGAACGCAAGTCCCTTGTGCTTACCGGTATCTGGTAG 2095
Db 2062 GGAACCATGTGCTCAGAAAGAAATTAATGCTAGGCTTAAACAGGCGAGCTGGTGG 2121
Qy 2096 ATTCACCACTTTCAGAGTTTGGTTCAGAGAGTCTCAAGACAGACAGACCTCA 2155
Db 2122 ATCTCTCTTCCATGAATTTGCTACGAGAGTCTGAGAACGTCGCGCGCTCTTACC 2181
Qy 2156 AAGTCCCAACCATGCGAGTCTATGAGAGTCCAGGTTCAAGTAAATCTGGAATCA 2215
Db 2182 AAGTACCAACCATAGGGGTGTATGGCTGTCCGGGGTCAAGCAAGTCTGGCATTTAAA 2241
Qy 2216 GCGCTGTGACTTAAGAAAGATCTGTTGAGTGGAGAGAGAAATCTGGCGAATCA 2275
Db 2242 GCGGAGTCAACCAAAAGATCTGGTGTGAGCGCCAGAAAGAAACCTGCCAGAAATA 2301
Qy 2276 TCAGAGATGTAGAGAGATGAGAGTATGATGTTGCTAGGACTGTGCATTCAGTGC 2335
Db 2302 TAAAGGAGTCAAGAAATGAAAGGGCTGAGCGTCAATGCAAGACTGTGACTAGTGC 2361
Qy 2336 TTCTTAATGGGTTTAAGCAACCCGTTAACTCTGTACTTATGATGAGCAATTTGCTGCC 2395
Db 2362 TCTTGAATGATGCAAAACCCCGTAAAGACCTGTATTTATGACGAACCTTTTGTGTTGC 2421
Qy 2396 ATGCAAGGACGCTGCTGCACTGATTTGCTCAATGCTCAAACTTAAGAAAGTGTATTTGCG 2455
Db 2422 ATGCAAGGACCTCTCAGAGCGCTCATAGCCATCATTAAGACTTAAAGGCAAGTGTCTGCG 2481
Qy 2456 GGGAGCCCAAAACAATGGCGCTTCTTTAATGATGTGCTGAAAGTAAATTTAACCTG 2515
Db 2482 GGGAGTCCAAAACAAGTGTGCTTTTCAATATGATGTGCTGAAAGTGAATTTTAAACCG 2541
Qy 2516 ACATATGCACTGAAGTGTACCAATAAAGCATCTCTAGAGGCTGACACAGACTGTAAACG 2575

Db 2542 AGATTTSCAGCAGAGTCTTCCACAAAAGCATCTCGCCGTTGCACATAATCCGTGACTT 2601
Qy 2576 CCATGCTTCCAGCTCTTTCTAAGCAAGGAAATGAAGAGGTTTAAACCATATGCTGATA 2635
Db 2602 CCGTGTCTCAACCTGTGTTTACGACAAAAGATGAGAACGAGAACCCCAAGAGACTA 2661
Qy 2636 AATCATCATATAGATACCAAGAGGACCAAGAGCGCACAAAGATGATCTGATTTAACT 2695
Db 2662 AGATTTGATTTGACACTACCTGCGCATGACCAACCGAAGAGAGAGCATCTCATCTCT 2721
Qy 2696 GTTTCAAGGATGGGTGAAAACGCTACAGATTGACTTCAAAAATCAGAAATCATGACTG 2755
Db 2722 GTTTCAAGGATGGGTGAAACAGTTGCAAAATAGTTTCAAAAGCCAAAGAAATATGACG 2781
Qy 2756 CCGCTGATCGCAAGACTTAACCGGAAAGCGTTTATGTGTCAAGTAAAGTCAAG 2815
Db 2782 CAGTGTCTCTCAAGGGCTACCCGTAAAGCGGTATGTGCGGTGACAAAGGTGAATG 2841
Qy 2816 AGAATCCACTCTACTCGCAGACTTCTGAGACGTAAGCTGTTACTTACAGCACAGAAA 2875
Db 2842 AAAATCCCTGTACGACACCACTCAGAACATGTGAACGTCTACTGACCCGACGAGG 2901
Qy 2876 AACGATTTCTGGAAGACGCTAGCTGTGATCTCTGATTAAGACACTTACAGCTTAAT 2935
Db 2902 ACCGTATCGTGTGAAAACACTAGCCGCTGATCATGATTAATAATCTGACGCGCAAGT 2961
Qy 2936 ATCCCGGGATTTTCAACGCTTCAATGAGACAGTGGCAGCCCGAAACAGAGCCATTTAG 2995
Db 2962 ATCTGGAACTTCACTGCGCACATAGAGAAATGGCAAGACATGAGTCAATCATGA 3021
Qy 2996 CACGCTTCTGTGATTAAGCCGACAGACGCTGATGTTCAGAAATAGAGTGAACGTCTCT 3055
Db 3022 GGCACATCTTGAAGAGACCGGACCTTACCGACGTTTCCAAAATTAAGCGGAAGTGT 3081
Qy 3056 GGGGAAAGCTTTAGAGCCAGTCTTGGCCACGGCCAACTTTGCTGACGAGACAGACT 3115
Db 3082 GGGCCAAAGCTTTGTGTCCGCTACTGAAGACTGACAGCATATGACATGACCATGMAAAT 3141
Qy 3116 GGGAGAGTGTGACCCCATTTAGAGATGACAGAGCTGATCTGACCTGAAGTGGACCTGAACT 3175
Db 3142 GGAACACTGTGATTACTTGAAGCGAACAAGTCACTGACGAGATATGATTAAGC 3201
Qy 3176 TCTTTGACACAGGTTCTTTGAGTGAAGCTGACAGTGGGTTATTTCCGCTCTACCG 3235
Db 3202 AACTATGCGTGAAGTTCTTTGAGTGAAGCTGAGCTGAGCTCGGCTATTTTCTGACACCACTG 3261
Qy 3236 TCGCACTTACTTACAGGATTCAGACTGGGATTAATCTGCCAAGGAGAACATGTATGGGC 3295
Db 3262 TTCCGTTATCCATTAGGAATTAATCACTGGGATTAATCCCGTGCCTTAACATGTACGGGT 3321
Qy 3296 TTAATAGAGTGTGCAAAAGAGTGTTCACAGCGCATTCCTGTGATCACAAAAGCGGTTG 3355
Db 3322 TGAATTAAGAGAGTGTCCGACGCTCCGCAAGTAAACCAACTCTCGAGCAATTTG 3381
Qy 3356 ACACAGCAGAGGTGCTGATTAAGATTAAGATTAATACATCAAGSACTACTCTCAACAATTA 3415
Db 3382 CCACCGAAAGTGTATGATGATGAACACTGGACGCTGCCCAATTAATGATCCGGCATTA 3441
Qy 3416 ATGTGTTCCATTAAATGCGCGGTTGCCCACTGCTTATGCTGTGACCAAGAGACAGG 3475
Db 3442 ATCTAGTACTGTGAACAGAAAGACTGCTCATGCTTTAGTCTTCCACATATATMAAC 3501
Qy 3476 GTACAACTGATCAAGGAGATTCCTATCTAAGATGAAGGGCAATCTGTGTGTGATG 3535
Db 3502 CACAGAGTGAATTTTCTTCAATTCGTACGAAATGAAAGGCGAGAACTGTCTTGTGTG 3561
Qy 3536 GCGATCTTATCAGCATTCAGAGGAAAGAAATAGAGTCAATGGGCTCATTTGCCCATATA 3595
Db 3562 GGGAGAACTGTCTGCTCCAGGCAAAAGGTCACTGTGTGTGACACAGCTGAGAGCTA 3621
Qy 3596 CCATCAGGTGTATCTGATTTGGGAATACCTAGCCATGTGCGTAAATATGACATTATCT 3655

Db 3622 CCTTAGAGCTCGGCTGATTTAGGTATCCAGGTGACGTCGCCAATAACGATTTGAT 3681
Qy 3656 TTGTCATAGTTAGAGCCCGGTACAGAAACCATCACTAACAGAGTGGAGATCAAGCTA 3715
Db 3682 TTATTAAGGTAGAGACTCATATTAATATCATATTTATCAGCAGTGTGAAGCCAGCA 3741
Qy 3716 TCCACACAGACATGCTAAAGTGTAAAGGCTGTCACACCTGAAACATGCGGAGATGTG 3775
Db 3742 TTAAGCTAGTATGTTAGCAAGAAAGCTGTCTGATTTGAATCCCGCGGAACTGTG 3801
Qy 3776 TGGCTATAGGATAGGGCTGTCTGATCCGCAACCGAATATCATCTGCGGTGAC 3835
Db 3802 TCAGCATAGTATGTTATACGCTGACAGGCGCAGAGCATCTATGCTATAGAGCGC 3861
Qy 3836 GCTCATATAGGTTTACCGGTGTCTGTACGCTTAAAGACATGCGAATAATAGAGGTTG 3895
Db 3862 GGCAGTTCAAGTTCTCCCGGATATCAACCAAAATCTCTACATAGAGAGAGAGATAC 3921
Qy 3896 TCTTGCTGTTCTTCGCAAGAGCAACCGGCACACACATGACAGAGACGAGTCCGCTG 3955
Db 3922 TGTGTATATGTTGAGTACGATTCGCAAGGCCGTATCGACATCTTACAGCTTTGAT 3981
Qy 3956 TAGTCTTGACACATCTATCAAGGTCAACCAAGTACGAGCGAGAGAGCTCCAGCT 4015
Db 3982 CTACCTTGAACCAACATCTATACAGGTTCAGACTCCAGCAACCGAGTCCGACCTCAT 4041
Qy 4016 ACAGATGATCAGAGGTGATTTAGCAGAGCGGTGACCAACTATGCTTATGCTGCTA 4075
Db 4042 ATCATGTGTGCGAGAGGATATGCGACCGCACGAGAGATGATCATAAATGCTGCTA 4101
Qy 4076 ATAGCAAGGTCAACCAAGTTCGAGATGTCGCGGTGACCTGATACCAAAATGCGCGCTG 4135
Db 4102 ACAGCAAGAGCAACTCTGCGGAGGGGTGTGCGAGCGCTGTATAGAAATTTCCGGA 4161
Qy 4136 CTTTGTATGACAGCCATAGCTGTGCGGACGCTAGACTGTGAACGACAC--GC 4192
Db 4162 GCTTCGATTTACGCGCATCGAAGTGAAGAAACCGCATGCTGTAAGGTGACGCTAAC 4221
Qy 4193 TCATCATCATGCTGAGAACCATTTTCTAAGATGCGGACCGGAGCGGACCTTA 4252
Db 4222 ATATATATTCATGCTGAGAACCAACTTCAACAAAGTTTGGAAATTTGAGGGGCAAC 4281
Qy 4253 AGCTGCACTGCTCTACATGAGCATAGGTCATGTCATGTCACGCTGAGCGATTAACAAA 4312
Db 4282 AGTTGGCAGAGGCTTATAGTTCATCGCTAAATTTGTCAACATTAACAATTAAGTCA 4341
Qy 4313 TATCATGACGCTGCTGACACGCGATCTATTTGCTGGCAAGATGAGTATGATCAAT 4372
Db 4342 TAGCATTCACATGTTCTCACCGGCATCTTTCCGGGAAACAAAGATGACTAACCAAT 4401
Qy 4373 CATTGCATCACCTGCTTCTGACACTAGCGGATGCGGATGTCACCAATATTTGCT 4432
Db 4402 CATTTAACCATTTGCTGACAGCTTTAGACACACATGATGCAATATGACTCA 4461
Qy 4433 TGGATTAACAATGAGAGACAGGATATGAGGCAATTCACCGCAAGAAAGAGCTGAAA 4492
Db 4462 GGGAAAGAAATGGAATGACTCTCAAGAGACAGTGGCTAGAGAGAGAGTGGAGG 4521
Qy 4493 TTCTGGA-----TGATGACAAAGCAGTACGACTTGTGCTGAGGCTC 4537
Db 4522 AGATATGATATCAGACACTTTCGTTGACAGAACCGGATGACAGAGCTGTGAGGTAC 4581
Qy 4558 ACCCAAGAGCTTTTGGCAGGCAACAGTTACTCCGTCAATGAGGCAAGTTGTATT 4597
Db 4582 ATCCAAAGATTTTGGCTGAGAAAGGCTACAGCAAGATGAGCAAGACTTTCT 4641
Qy 4598 CATACCTGAAAGTACAGATTCATCAGACCGCAAGGACATTCGCAAAATCATGCAA 4657
Db 4642 CATATTTGGAAGGACCAATTTCAACAGCGGCGCAAGATATGCAAAATTAATGCA 4701
Qy 4658 TGTGCGCCAAATCTGAGGCTTAATGAGCAGATTTGCTGTATCTCTGGGGAGAGTA 4717
Db 4702 TGTGCGCAGTTGCAACGAGGCGCATGAGCAAGTATGATATCTCTCGGTGAAGCA 4761

Qy 4718 TGTCCAGCATCGGCTCCAAATGCCCAGATGAGAGGTGAGAGCGCTGTGCTCACCTCA 4777
Db 4782 TGAGCAGATTAAGTTCGAAATGCCCCGTCGAGAGTCCGAAAGCTTCCACACCACTAGCA 4821
Qy 4778 CACTTCATGCTGTGTATTAACGTATAGCGGTGACGCGGTATACAGTTCGCTGTG 4837
Db 4822 CGCTGCTTGTGCTGTGATCATGCTATGACTCCGAAAGAGTACAGCGCTTAAGAGCT 4881
Qy 4838 CGAAGAAAGACAGTTCGCCGTATGCTATCATTTCTGTTGCCAAGTACAGATACAG 4897
Db 4882 CACGTCGAAACAAATATCTGTGCTCATCTTTTCCATTCGCAAGTATGAAACACTG 4941
Qy 4898 GCGTCAGAGCTACAGTACGAGCAACAGCTGCTTTTCCAGGCGCTGTACACCGGCTG 4957
Db 4942 GTGTGCAAGAAATCAGTGTCTCCAGCTTATCTGTTCTCACCGAAGGTGCTGCTACA 5001
Qy 4958 TACACCCAGAGATGAC 4974
Db 5002 TTCAATCCAGGAGATGAC 5018

Db 5002 TTCAATCCAGGAGATGAC 5018

RESULT 10
VEUS5347
LOCUS
DEFINITION
Venezuelan equine encephalitis virus strain 6119 nonstructural
polypeptide and structural polypeptide genes, complete cds.
ACCESSION
U55347
VERSION
U55347.2 GI:20800451
KEYWORDS
ORGANISM
Venezuelan equine encephalitis virus
Venezuelan equine encephalitis virus
Viruses, ssRNA positive-strand viruses, no DNA stage; Togaviridae;
Alphavirus; VEEV complex.
REFERENCE
1 (bases 1 to 11423)
Weaver, S.C., Salas, R., Rico-Hesse, R., Ludwig, G.V., Oberste, M.S.,
Bohnel, J., and Tesh, R.B.
Re-emergence of epidemic Venezuelan equine encephalomyelitis in
South America. VBE Study Group
Lancet 348 (9025), 436-440 (1996)
JOURNAL
MEDLINE
PUBMED
8709783
2 (bases 1 to 11423)
Weaver, S.C.
Direct Submission
Submitted (18-APR-1996) Scott C. Weaver, Pathology, Univ. TX Med.
Branch, 301 University Ave., Galveston, TX 77555-0609, USA
JOURNAL
TITLE
REFERENCE
AUTHORS
Weaver, S.C.
Direct Submission
Submitted (01-MAY-2001) Scott C. Weaver, Pathology, Univ. TX Med.
Branch, 301 University Ave., Galveston, TX 77555-0609, USA
REMARK
COMMENT
On May 15, 2002 this sequence version replaced gi:2335117.
FEATURES
source
1..11423
/organism="Venezuelan equine encephalitis virus"
/mol_type="genomic RNA"
/strain="6119"
/db_xref="taxon:11036"
/note="antigenic IC"
25..7506
/note="contains low efficiency read through at stop codon;
encodes nsP1, nsP2, nsP3 and nsP4 proteins"
/codon_start=1
/transl_except="(pos:5662..5664,aa:OTHER)
/product="nonstructural polypeptide"
/protein_id="AA028637.1"
/db_xref="GI:20800451"
/translation="MERKVDHIDEDSPFLRLQSPQFEVAKQVTDNDHANRAPS
HLSKLIETEDVPSDITLIDGAPARMYSGRKHHCIPKNCADPRLVYATRLK
NKEITDKELDKKEELAAVMSDDLEETMCLHDDSCRYEGQVAVYQDVAVDGP
SLYQANKGVAVAVIGPDTTPFPFKULAGAYPEYSINWMDVETLFRANIGLSSDVM

Db 1022 ATGTCCAGCTACATTTGTGTGAACAAATGACTGCACTGTGGCAACAGATGTCAGTGCGG 1081
Qy 1076 ATGAGCACAACAACTATTTGGGCTCAACCAAGATGTGCTCATTTGATAGGAGCG 1135
Db 1082 AGGAGCGCCAAAATCTGCTGTTGGCTTACCAAGGCGATAGTCGTCAACGGTCCGACCC 1141
Qy 1136 AAAAATACTAACAATGACAACTATCTATTACAGTGTGCGCCAGCGTTTTC 1195
Db 1142 AAAAAGAACCAATACATGAGAAATATCTTTTGGCCGTAGTGGCCAGGCAATTTGCT 1201
Qy 1196 GGTGGCGCGTGAACATCGTCCGACTTGGACGACGAGAAAGAACTAGGGGTCCGGAGC 1255
Db 1202 GGTGGCAAGAAATATAAGAAACCAAGAAATGAGGCGCAGTACGACTACAGATA 1261
Qy 1256 GCACTCTTATCTATGGGCTGTGCGGCTTTCAGACCCAGAAATTCATGCTCATCTCA 1315
Db 1262 GACAGTTAGTCAATGGGCTGTGCTGCTTTTGAAGGCAAGATTAATCTATTATTA 1321
Qy 1316 AGAAGCTGTGTACGCAACATTTAAGAAAGTACTGCGCTTTTGTGACTTGTGATTC 1375
Db 1322 AGCGCCAGATACCAAAACATCATCAAAAGTAAAGCGATTTCCATCTATTCGTGCTGC 1381
Qy 1376 CAGCGCTTACCGCAACGCGGCTGATATGGGCTTCCGCGTGAAGCTCAAGCTGCTTG 1435
Db 1382 CCAGATATGGCAGTAAACACTGAGATGCGGCTGAAACGAAATCAAGAAATGCTAG 1441
Qy 1436 AACCAATGTCAACCCGACCGGCTATTACAATGGCCGATGGAGCATCTGCGTGGCT 1495
Db 1442 AAGACACAAGAGCGGTCACTCTCATTAATGCGAGACATACAAAGGCTTAAGTGG 1501
Qy 1496 TACACCAAGAACTGAAGAAAGTGTGACGCGGAAGAGATCAGAGAGCCCTGCCACCT 1555
Db 1502 CAGCGGATGAGGCTTAAGAAAGTGTGAGAGCGAGAGAGTGGCGCTGACATACACCTT 1561
Qy 1556 TGCTCCCTGAATTAAGAAAGAGCCGTAGAGGCGAGAGTGAAGCTTATATGCAAGAG 1615
Db 1562 TGGCAGTGTATTTTGAAGAGCCACTCTGGAAGCCGATGTCACTTGAATGTTTAAAGG 1621
Qy 1616 CAGGACAGGTATGCGTGAAGACACCAAGAGACATCGAGGTGAACAATTTCCAGAGG 1675
Db 1622 CTGGGCGCGCTCAAGTGAAGACCTCGTGGCTTGAATTAAGGTTCACAGCTATGCGGCG 1681
Qy 1676 AAGAGAAATTTGGCTTACGCTATATCTTCAACCCAGCGGCTATTTGAATAGTAAAAAC 1735
Db 1682 AGGACAAATGTGGCTTACGAGAGTCTTCTCCAGAGCTGTATCTCAAGATGAGAAAC 1741
Qy 1736 TGGCGTATCAACCCATTTGGCGAAACAAGTACTGTATTAATGACTCAAAAGGTAGGCGAG 1795
Db 1742 TATCTTGAATTCACCTCTCGCTGAACAAGTATAGTATTAACAACCTGCGCGAAAG 1801
Qy 1796 GGAAGTACAAATCGAGGCTATACAGGTAAGTCAATTTACCAAGAGGAGCGCGGTCC 1855
Db 1802 GCGGTTATGCGCTGAACCTTACATGAAAGTGTGTGCAAGGAGCAATGCAATAC 1861
Qy 1856 CTGTTCAAGACTTCCAGGCTATGAGAGCGCTACGATCGTTTCAACGAGAGGAGT 1915
Db 1862 CCGTCAAGAGCTTTCAGAGCTCTGAGTGAAGTGCACATCTGTATCAACGAGAGT 1921
Qy 1916 TCGTAAACAGATATCTGCAACATCGCAATCAACGAGAGCGCTTAAACACTGAGAG 1975
Db 1922 TCGTAAACAGATATCTGCAACATCTGCAACATGAGAGCGCTTAAACAGATGAG 1981
Qy 1976 AGTACTATTAAGCTGTAAAGACTCAAGACACAGACTCAGAAATACGTTCTGATATTGAGC 2035
Db 1982 AATATTACAAACCTGTCAAGCCAGAGACACAGCGCGAATACCTGTATGACATCGACA 2041
Qy 2036 CACGAAGTGTATTAGCAGAAAGACGAGGTCCCTTGTGCTTAAACGCGTGAATCTGTAG 2095
Db 2042 GGAACATATGCTCAAGAAAGATTAATGACTGTGGGTAGGCTTACAGCGAGCTGTGG 2101
Qy 2096 ATTCACCATTTACAGAGTTTGGCTACAGAGTCTCAAGACACGACGACGACCTCACA 2155
Db 2102 ATCTCTCCCTTCATGAATTTGCTAGAGAGTCTGAGAACGTCGCGCGCTCTTACC 2161

Qy 2156 AAGTCCCAACCATCGGAGTCTATGAGTGGCCAGGTTCAAGTAAATCTGGAATCATCAAA 2215
Db 2162 AAGTACCAACCATGAGGCTGTATGGCTGTGCGGGTCAAGCAAGTCTGGCATATTAAA 2221
Qy 2216 GCGCTGTGATTAAGAAAGATCTGTTGTGAGTGCAGAAAGGAAAATCTGCCGAAATCA 2275
Db 2222 GCGCAGTCAACCAAAAGATCTGTTGTGAGGCGCAAGAAAGAAAATCTGCCGAAATTA 2281
Qy 2276 TCAAGGATGTAAAGAGATGAGACGTATGATGTGTCTGTATGAGCTGTCAATTCAATGC 2335
Db 2282 TAAAGGAGCTCAAGAAATGAAGAGGCTGAGACGTAAATGCCAGAACTGTGACTAGTGC 2341
Qy 2336 TTCTAAATGGGCTTAAAGACCCGTTTAAACCTGTATGATGATGAGGATTTGCTGCG 2395
Db 2342 TCTTAAATGATGCAAAACCCCGTATGAGACCTGTATATTTGACAAAGCTTTTGTGTC 2401
Qy 2396 ATGCAAGGACGCTGTGCACTGATTTGCCATGCTCAAACTTAAGAAAGTGTATTGTGCG 2455
Db 2402 ATGCAAGGACCTCTCAGAGCGCTCATAGCCATCATTAAGACCTTAAAGGAGTGTCTGCG 2461
Qy 2456 GGGACCAAAACAAATGCGGCTTCTTAAATGATGTGCTGAAAGTAACTATTTAACATG 2515
Db 2462 GGGATCAAAACAGTGTGCTTTTCAATATGATGTGCTGAAGTGAATTTTAAACACAG 2521
Qy 2516 ACATATGACCTGAAGTATACCAATAAAGCATCTAGAGAGGTGCACAGACTGTACCG 2575
Db 2522 AGATTTGACCGAGGCTTTCCAAAGACATCTCTGCTGTTCCTAAATTCGTGACTT 2581
Qy 2576 CCATGCTCTCAACGCTCTTCTTACGACAAAGCAATGAAGACGTTAACCATGTGCTGATA 2635
Db 2582 CGGTGCTCAACCTGTTTATGACAAAGATGAAGACGACGAAACCGAAAGAGACTA 2641
Qy 2636 AATATCATATGATATCCACAGGACCAACAAAGCCGCAAAAGATATGATTTTAACTT 2695
Db 2642 AGATTGAATGACCTCTGCACTATACCAACCAAGACAGACATCTCATTTCTACCTT 2701
Qy 2696 GTTTCAGAGATGGGTGAACACGCTACGATTTGACTTAAACAAATACGAAATCATGACTG 2755
Db 2702 GTTTCAGAGATGGGTGAACAGCTTGAATTAAGTAAACGAAACGAAATTAAGACG 2761
Qy 2756 CGGCTGATTCGCAAGGACTTACGCGGAAAGCGTTTATGCTGTCAAGTAAAGTCAACG 2815
Db 2762 CAGCTGCTCTCAAGGCGTGAACCGTAAAGGCTGTATGCGGTTTGTGATCAAGGTAAG 2821
Qy 2816 AGATTCACCTTACTGCAAGCTTGTGACACGTGAACGTGTACTTACACGACAGAAA 2875
Db 2822 AAAATCCCTGTACCAACCACTCAGAACATGTGAACGTCTTACTGACCCGACGAGG 2881
Qy 2876 AACGATGTCTGGAAGACGCTAGCTGTGATCCCTGATTAAGACATTAACGATTAAT 2935
Db 2882 ACCGATGTGTGGAACCACTAGCGGTGATCCATGATTAACCACTGAGGCGCAAGT 2941
Qy 2936 ATCCGCGGATTTTCAAGGCTTCAATTTGACGACTGCAACGCGCAACGACGCTATG 2995
Db 2942 ATCTGTGGAACCTTCACTGCAAGATGAGAGATGSCAAGCAGAGATATGCAATCA 3001
Qy 2996 CAGCGTTTGTATTAAGCTGCAAGACGCTGATGTGTCGAAATTAAGTGAACGTCTGCT 3055
Db 3002 GGCACATCTTGAAGAGACCGACCTTACCGAGCTTTTCCAAAATTAAGCGAAAGTGTGT 3061
Qy 3056 GGGCGAAGGCTTTGAGCACTGTTGGCCACGCGCAACATTTGCTGACGAGACGAGT 3115
Db 3062 GGGCCAAAGCTTTGTGCGGATCTGAAGACTGCAAGCAATTAAGCAATGACATGAACAT 3121
Qy 3116 GGGAGACGTTGACCATTTCAAGCATGACAGAGCTTACTCACTGAAATGCGACTGAAC 3175
Db 3122 GGAACATGTTGATTTACTTGAAGCAAGCAAGCTCATCAGCAGAGATATGATTTGAAC 3181
Qy 3176 TCTTTGACCAAGGTTCTTTGAGTGAACCTGAGACATGGGTTATTTTCCGCTCTACCG 3235
Db 3182 AACTATGCTGAGGTTCTTTGAGCTGACCTGAGACTCGGCTATTTTCTCAGCACCACTG 3241

3236 TCGCACTTACTTACAGGATCAGCCTGGATTAATCTCGCCAGGGAAGAACATGATGGCC 3295
3242 TTCCGTTATTCATTAGGAATATATCACTGGATTAATTTCCCGCTCCGCTTACATGACGGGT 3301
3296 TTAATAGAGAGTACCAAGAGTGTCTACCGGCATATCCGTGATCACAAAAGCGGTG 3355
3302 TGAATAAAGAGTGTCCGACGCTCCCGCAGGTACCCACAATGCTGCTGACAGAGTTG 3361
3356 ACACAGGAGGAGTGTGATTAAGAAATATACCATCAAGAGTACTCTTCAACATTA 3415
3362 CCACCGAAGAGTGTATGACATGAAATCTGGCAGCTGCGCAATTAATGATCGCGCATTA 3421
3416 ATGGTGTTCATTAATCGCGGTGCGCCACTCGTTGATCGTTGACCAAAAGAGCAGG 3475
3422 ATCTAGTACTGTGAACAGAAAGTCTGCTCATGCTTTAGTCTTCAACCATATGACACC 3481
3476 GTACAACTGATCAGAGCGGATTCCTATCTAAGATGAAGGCGAAATCTGTGTGTGATCG 3535
3482 CACAGAGTACTTTCTTCAATTCGTCAGCAAACTGAAGGGCAGAACTGTGTGTGATCG 3541
3536 GCGATCTTATGACATTCAGAGGAAGAAAGTACGATGCGATGCTCATGCTCCACTATA 3595
3542 GGGAGAAAGTGTCCGTCACGCAAAAGGTGCACTGTTGTCAAGCCAGCTTAGGCTTA 3601
3596 CCATCAGGTGATCTGATTTGGGAATACCTAGCATGTGCGTAATATGACATTAATCT 3655
3602 CTTTGAAGTCTGCTGATTTAGTATCCAGGTGACGTCCCAATATGACATTTGAT 3661
3656 TTGTCAATGTTAGAACCCCGTACAGAACCATCACTACCAACAGTGCAGAGTACAGCTTA 3715
3662 TTATTAAGTGAAGCTCATATTAATACATCAATTAATGACAGTGTGAAGACACCGCA 3721
3716 TCCACACAGATGTAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 3775
3722 TTAAGCTTATGATGTGACCAAGAAAGTGTGATGATGATGATGATGATGATGATGATG 3781
3776 TGGCTATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3835
3782 TCAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3841
3836 GCTCATTTAGGTTTACCGGTGTCTGATGATGATGATGATGATGATGATGATGATGATG 3895
3842 GCGAGTTCAAGTTTCCCGGGTATGCAAAACCAATCTTCAATGAAGAGACAGAGTAC 3901
3896 TCTTGTGTGTTTTCGCGAGAGACACGCGAACACACATGACAGAGACAGACTCGGTG 3955
3902 TGTGTGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3961
3956 TAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4015
3962 CTACCTTACCAATCTATACAGTTTCAAGTTTCAAGTTTCAAGTTTCAAGTTTCAAGTT 4021
4016 ACAGAGTATCAGAGTGTATGACAGAGCGGTGACCAAGCTATGCTTATGATGATGATG 4075
4022 ATCATGTGTGTCGAGGGGATATGTCACAGGTCACGAGAGAGTATGATGATGATGATG 4081
4076 ATACCAAGAGTACACAGGTTCCGAGATGTGCGGTGATGATGATGATGATGATGATGATG 4135
4082 ACACCAAGAGTACACAGGTTCCGAGATGTGCGGTGATGATGATGATGATGATGATGATG 4141
4136 CTTTGTATGACAGCAATGCTGTGCGGACGCTAGACTGTGATGATGATGATGATGATGATG 4192
4142 GCTTCGATTTTACAGCCATGCAAGTGGAAAAAGCGCATGCTCAAGGTGATGATGATGATG 4201
4193 TCATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4252
4202 AATATATATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4261
4253 AGCTGCGAGCTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4312
4262 AGTTGGCAGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4321
4313 TATAGTACCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4372

4322 TAGGATTTCACTGTTGTCCACCGGATCTTTTCCGGAAACAAAGATGACTAACCCAT 4381
4373 CATTGATCACTCTGTTTACGCTTTTCACTACGATGATGATGATGATGATGATGATGATG 4432
4382 CATTGAACCAATTTGCTGACAGCTTTTACACCACTGATGATGATGATGATGATGATGATG 4441
4433 TGAATTAACATGGGAGACAGGATATGAGGCTATTCACCGCAAGAAAGCGTCGAA 4492
4442 GGGCAAGAAATGGGAATATGATCTTCAAGAGAGATGCTGATGAGAGACAGTGAAG 4501
4493 TTCTGGA-----TGATGACAGCCAGTGAATTTGATGATGATGATGATGATGATGATG 4537
4502 AGATATGATATATGACAGACTCTTTCGTCAGAGACCGGATGACAGACTGATGAGGATC 4561
4538 ACCCAACAGCTCTTGGCAGGACAGACAGGATATCTCGTCATGATGAGGCAATGATTA 4597
4562 ATCCGAAGATTTCTTGGCTGGAAGGAGGCTACACAGCAAGTATGATGATGATGATGATG 4621
4598 CATACCTGGAAGGATGACGATTCATGACACCGCAAGGACATTTGCCAATTCATGCA 4657
4622 CATATTTGGAAGGACCAATTTTCCACGAGCGGCAAGATATGACAGAAATTAATGCA 4681
4658 TGTGCCCCCAAAATCTGAGGCTTAATGACAGATTTGCTGTATCATCTGCGGAGAGTA 4717
4682 TGTGCGCAGTTGCAAGAGGAGCCATGACAGATGATGATGATGATGATGATGATGATGATG 4741
4718 TGTTCAGATCTCGCTCCAAATGCCCAGTACAGAGTACAGAGGCTGTCTTCACTACA 4777
4742 TGAGCAGACATTAAGTCCAAATGCCCCGTCAGAGAGTGTGAAAGCTCCACACACCTTGA 4801
4778 CACTTCATGCTGTGTATTAAGCTATGACGAGTGTGAGCGCGATACAGGTTGCGCTGTG 4837
4802 CGCTGCTTGTGTGTGATCATGCTATGATGATGATGATGATGATGATGATGATGATGATG 4861
4838 CGAAGAAAGAACAGTTGCGGTATGCTATCATCTTCTGTTGCCAATGACAGATGACAG 4897
4862 CAGTCCAGAAACAAATTAAGTGTGCTCATCTTTCATTTGCGAATGATGATGATGATGATG 4921
4898 GCGTGCAGAAAGCTACAGTGTGACAGAACCAAGTCTGTTTTCAGGCGTGTACACCGGCTG 4957
4922 GTGTGCAGAAAGATTCAGTGTCTCCAGCTTATGATGATGATGATGATGATGATGATGATG 4981
4958 TACACCCAGAGATGAC 4974
4982 TTCAATCCAGAGATGAC 4998

RESULT 11
EVNSPERA 11447 bp ss-RNA linear VRL 25-MAY-1999
DEFINITION Venezuelan equine encephalitis virus, complete genome.
ACCESSION L01442
VERSION L01442.2 GI:4887231
KEYWORDS
SOURCE
ORGANISM
Venezuelan equine encephalitis virus
Venezuelan equine encephalitis virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
Alphavirus; VEEV complex.
REFERENCE
1 (bases 1 to 11447)
Kinney,R.M., Johnson,B.J., Brown,V.L. and Trent,D.W.
Nucleotide sequence of the 26 S mRNA of the virulent Trinidad
donkey strain of Venezuelan equine encephalitis virus and deduced
sequence of the encoded structural proteins
Virology 152 (2), 400-413 (1986)
JOURNAL
MEDLINE
PUBMED
3088830
REFERENCE
2 (bases 1 to 11447)
Johnson,B.J., Kinney,R.M., Kost,C.L. and Trent,D.W.
Molecular determinants of alphavirus neuroinulence: nucleotide and
deduced protein sequence changes during attenuation of Venezuelan
equine encephalitis virus
J. Gen. Virol. 67 (Pt 9), 1951-1960 (1986)

[illegible][illegible]

old_sequence 106..107
/gene="NS"
/note="this correction changes amino acid nsP1-21-Thr to nsP1-21-Ser, this 1992 data corrects the 1989 reference in citation [3]"
/citation=[3]
/replace="gc"
1533..1534
/gene="NS"
/note="this correction results in changing amino acid nsP1-497-Arg to nsP1-497-Ala, this 1992 data corrects the sequence reported in the 1989 reference in citation [3]"
/citation=[3]
/replace="gc"
1696
old_sequence
/gene="NS"
/note="this correction results in changing amino acid nsP2-16-Asp to nsP2-16-Ala, this correction adds another difference between the Trinidad donkey and vaccine TC-83 strains reported in citation [3]"
/citation=[3]
/replace="c"
7532..11444
/gene="26S mRNA"
7532..11444
/gene="26S mRNA"
/note="the structural proteins of the virus are translated as a polypeptide precursor molecule from an intracellular 26S mRNA species that is identical to the 3'-one-third portion of the viral genomic mRNA; putative"
/function="intracellular, subgenomic viral mRNA species"
Query Match 42.4%; Score 2119.6; DB 14; Length 11447;
Beet Local Similarity 64.6%; Pred. No. 0; Mismatches 1744; Indels 24; Gaps 3;
Matches 3220; Conservative 0;
gene
mRNA
10 ACTATGATCCATATGGAAGAATTACGTTGACTAGATGTCAGAGCCCGATGTC 69
30 AATTACTACCAAAATGGAAGTTGACCTTGACATCGAAGAACAGCCCATTCCTC 89
70 AAGTCGTTACAGCGGACGTTTCCAAATTTGAGATCGAAGCAGGCACTGACAT 129
90 AGAGCTTTGACGCGAGCTTCCGCACTTGAGTGAAGCCAGCAGTCACTGATAT 149
130 GACCATGCAATGCGCAGCGCTTTCCGATGTCGGAACAAAGCTCATTTGAGAGCA 189
150 GACCATGCTAATGCGAGCGTTTTCGATCTGGCTTCAAACTGATGAAACGAGGTG 209
190 GACCGGACCAAGTATCTTGAACATTTGGAAGTGGCCGCTCAGACATGACATTCAT 249
210 GACCATCCGACAGATCTTGACATTTGGAAGTGGCCGCGCCGCAAGATGATTTCT 269
250 CACCGTATCATTTGATCTGCTTATGATTAAGCGCTGAAGACCCGCGACAGCTCA 309
270 CACAAGTATCATTTGATCTGCTGATGATGATGTCGGAAGATTCGGAAGATTTGA 329
310 TATGCAAGAAACTTAAGAAAAGT-----GACATTTACCGACAGAAATAGCTCT 363
330 TATGCAACTAAGCTGAAAGAAAAGTAAAGAAATTAAGTAAAGAAATTTGAACA 389
364 GCGGACAGCTGCTGGAAGTATGTCACACAGAGCAGAGACTCCATCTCTGTGTATG 423
390 ATGAAGAGTCTGCGCCGCTCATGAGGACCTTGAACCTGGAACCTGATATGTGCTC 449
424 CACACAGACGCGACGCTGATGATCTTTGAAAGTGAAGATTAACAAGATGTGACGA 483
450 CACGACGACGAGTGTGCTGCTACGAGAGGCGCAAGTGGCTTTTACCAAGATGTA 509
484 GTCCATGACCGACATCATTTACCAAGAGCGCTTAAAGAGTGAAGCAATTTACTCG 543
510 GTTGACGACCGACAGATCTCTATCAACCAAGCCATTAAGGAGTTAGTGTGCTTCTG 569
544 ATAGGCTTTGACAGACCCCTTTATGTATCAAAAACATGGCAGGTTCCTACCTACTAC 603

Db 570 ATAGGCTTTGACAGACCCCTTTTATGTTTAAAGACTGGCTGAGATATCATCTAC 629
Qy 604 AACACGACTGGGCTGACGAGAGATTTGAAAGCAGCTAACCTTGGCTTCGTA 663
Db 630 TCTACCAACTGGGCGCGAAGCCGCTTAAACGCTCTGTAACATTAAGCCCTAT 689
Qy 664 GATCTTGAAGAGCAGGCTTGGAAAACCTCAATCCTTGAAGAAAGAGCTCAACT 723
Db 690 GACCTTATGAGCGGCTGACGTAAGGAGATTCATTTTGAAGAAAGTATTTGA 749
Qy 724 ACTAATAGATCATATTTCTGCTGTTTCAACAATCTACAGAAAGATAGTCACT 783
Db 750 TCCAAAGATTTCTATCTCTGTTGGCTGACCATCTACACAGAGAAAGAGCT 809
Qy 784 GCTAGCTGCACTTCCAAACGTTTCCATCTTGAAGAAAGTCACTTCAAGATGA 843
Db 810 AGGAGCTGGACCTGCGCTGTATTTCACTTACGTGGGCAAGCAAAATTAACA 869
Qy 844 TGTGGGACATTTGACGCTGTAAGGAGTACATCAAAAAGATTAAGATCAGCC 903
Db 870 TGTGAGCTATATGTTAGTTGTCAGCGGTACGTTTAAAGAAATGCTATCAG 929
Qy 904 CTATACCGTAAAGTTGAGAACTTGGCGTCCCAATGATCGAGGCTTCTTGA 963
Db 930 CTGATGGAAGCCTTCAAGCTATGCTGCTACATGATGACCGGAGGATTTCT 989
Qy 964 AAGTCAAGATACGCTGCGCGCGAGAGGTTCTTTCTGTGTATGATATCA 1023
Db 990 AAGGTGCAACACATTTGAACGCGGAGAGGCTCTTTTCCGTGTACAGTATG 1049
Qy 1024 GCCACACTTGGCATGATGACAGAGGATCTGGCAATCTAGCTAGTGTGATGAC 1083
Db 1050 GCTATGTTGTGACCAAAATGATCGGATCTGGCAACATGTCACTGCGGAGAC 1109
Qy 1084 CAAAACTATTTGTTGGCTCAACCAAGATTTGCTCAATGATGATGAGACGAA 1143
Db 1110 CAAAACTGCTGTTGGCTCAACCAAGATTTGCTCAACGCTGCGACCAAGAA 1169
Qy 1144 ACTAACAATGCAAGATCTATTTACAGTGTGCGCCGAGCTTTTCAAGTGG 1203
Db 1170 ACCAATACATGAAATTTACCTTTTCCGTGTGTCGAGGATTTGATGTT 1229
Qy 1204 CGTGAATGTCGCGCTTGAACGAGAGAAAGATTAAGGAGGCGGAGCGCACT 1263
Db 1230 AAGGAATTAAGGAAGTCAAGAAATGAAAGGCTAGAGCTACAGATTAAG 1289
Qy 1264 ACTATGGCTGCTGCTGGGCTTCAAGACCCAGAAATCAATCATCTCAAGAG 1323
Db 1290 GTATGGGCTGTTGTTGGCTTTAAGAGCAAGATTAATCTATTTAAGCGCC 1349
Qy 1324 GGTACGAACATTAAGAAAGTACCTGCTCTTATCATCTTTGATTTCCAGCT 1383
Db 1350 GATACCCAAACATCACTAAAGTGAACAGGATTTCCATCTATTTGCTGCC 1409
Qy 1384 ACCAGCAGCGGCTGATATGAGGCTTCCGCTAGCTCAAGCTGCTTGAACA 1443
Db 1410 GGCATGATCACTTGGATGCTGCTGAGAACAAAGATCAGAAATGTTAGAG 1469
Qy 1444 GTCAAAACCCGCTGCTATTTAATGATGCGATGTGAGCATCTGCTTACAG 1503
Db 1470 AAGGAGCGTACCTCTCATTTACCGCGAGAGCTACAAAGATTAAGTGG 1529
Qy 1504 GAACTGAAGAGTGTGCTGACGCGGAAAGATCAGAGAACCTTCCACTTCT 1563
Db 1530 GAGCTTAAGAGGCTGCTGAAGCGAGAGTGTGCGCACTTCAACCTTTGG 1589
Qy 1564 GAAATTAAGAAAGACGTAAGGAGAGATGACCTCAATTTGCAAGAGCGAG 1623
Db 1590 GATGTTAGAGGCCCATCTTGAAGCGGATGTCACTTATTTACAAAGGCTGG 1649
Qy 1624 GTAGCGTGAAGACCAACAGAGACATCAAGGTGCAAGTTACCAAGGCGAG 1683

Dh 1650 GGCTCAGTGAGACACTCGTGGCTTGAATTAAGCTACCGCTGCGCAGACAAAG 1709
Qy 1684 ATTGGGCTTACGCTATACCTTTCACCCGAGCGGATTGAATGTAAGAAATCTGGCGTGT 1743
Db 1710 ATCGGCTCTTACGCTGTGCTTCTCGCAGGCTGTACTCAGAGTGAATTAATCTTGC 1769
Qy 1744 ATCCACCATTTGGCGAAACAAGTACTGTATGACTCAMAAGTAAAGGCGAGAGATAC 1803
Db 1770 ATCCACCTCTCTGTCGAACAAGTCAATGATTAACACTCTGCGCAAAAGGGGCTTAT 1829
Qy 1804 AAAGTCAGCCCTTACCACGGTAAAGTCAATTTGACCAAGAGGACGGCGCTTGTCAA 1863
Db 1830 GCCGTGGAACCAATACATGTGTAAAGTGTGCGCAAGGGAACATGCAATACCCCTCCAG 1889
Qy 1884 GACCTTCAAGGCAATTGATGAGAGCGCTACGATGTTTTCAAGAGAGGAGTTGTTAAAC 1923
Db 1890 GACCTTCAAGCTCTGATGTAAGTGTCCACCATTTGTGTACAGAACGTGAGTTCGTAAAC 1949
Qy 1924 AGATACCTGCAACACATCGCAATCAACGAGAGCGCTAAACATCTAGCGAAGTACTAT 1983
Db 1950 AGGTACTGTCAACATATTGCCACATGAGGAGCGGTGAACATGATGAATATTTAC 2009
Qy 1984 AAGATCTTAAGACTTCAGACACACATCTCAGATACGTTCTTGATTTGACGCAAGAAAG 2043
Db 2010 AAAATCTCAAGCCGAGGAGCAGACGCGCAATACCTGTACGACATCGACAGAAAGAG 2069
Qy 2044 TGTGTTAAGCGAGAGACGACGAGTCCCTTGTGCTCAACCGGTGATCGTGTGATCCACA 2103
Db 2070 TCGCTCAAGAAAGAACTAGTCACTGAGGCTCAGAGGCACTGAGTGTGATCTTCCC 2129
Qy 2104 TTTCAAGATTGCTGACAGAGTCTCAAGACACAGCAGACACTCACAAGTCCCA 2163
Db 2130 TTCCATGAAATTCGCTTACGAGAGTCTGAGAACACACAGCGCTCTTACCAATACCA 2189
Qy 2164 ACCATCGAGTCTATGAGTGTCCAGGTTGAGTAAATCTGAAATCATCAAAAGCGCTGTG 2223
Db 2190 ACCAATGAGGAGTAAAGGCGTGCAGAGATCAGCAAGTCTGATCAATTAAGCCGCAATC 2249
Qy 2224 ACTAAGAAAGATCTGTTGTGTGAGTGCAGAAAGAACTGGCGAAGATCATCGAGAT 2283
Db 2250 ACCAAAAAGATCTAGTGTGAGCGCCCAAGAAAGAACTGTGCAAGAAATTAAGGAGAC 2309
Qy 2284 GTAAGAGATGAGACGTATGATGTGCTGCTAGAGCTGTGATTCAGTCTTAAAT 2343
Db 2310 GTCAAGAAATGAAGAGGCTGAGCTCAATGCAAGACATGAGATCTGATGCTTGAAT 2369
Qy 2344 GGGGTTAAGCACCCCGTTAACTCTGTACATTTGATGAGGCAATTTGCTGCCAGCAGG 2403
Db 2370 GGATGCAAAACACCCCGTAGAGACCTGTATATTGACGAAGCTTTTGTGTGATCAGAGT 2429
Qy 2404 ACGCTGTGCACTGATTTGCCATCTCTCAAACTTAAGAAAGTGTATTTGTGGGGGACCA 2463
Db 2430 ACTCTCAGAGGCGCTATAGCCATTATTAAGCCTTAAGAGGAGTCTCTGCGGGGATCCC 2489
Qy 2464 AAACATACGCGCTTCTTAAACATGATGTGCTGAAGTACATTTTAACATGACATATGC 2523
Db 2490 AAACAGTGGCTTTTAACTATGATGTGCTGAAGTGTCAATTTTAACACAGAGATTTGC 2549
Qy 2524 ACTGAAGTACCATTAAGACATCTCTAGAGAGTGCACACAGATGTAAACGCCATCTGC 2583
Db 2550 ACACAAAGTCTTCAACAAAGACATCTCTGCGGTTCACTAAATCTGTGCTTCCGTCGTC 2609
Qy 2584 TCCAGGCTCTTCTACGACAGAGATGAAGCGGTTAAACCATGTGCTGATTAATCATC 2643
Db 2610 TCACCTGTTTAAAGCAAAATAATGAAGAGATCCGAAAGAGACATTAAGTTGTG 2669
Qy 2644 ATAGATACACAGGACCAAAAGCGCACAAAGATGATCTGATCTTAACCTGTTTCA 2703
Db 2670 ATTGACATCACCGAGATACCAAACTTAAGCAGAGCGATCTTATCTCATTTGTTTCA 2729
Qy 2704 GGATGGGTGAACAGCTACAGATTGATCAAAATAATCAAGAAATGATGATCGGCTGCA 2763
Db 2730 GGGTGGTGAAGCAATGATTAATTAAGCAAGCAAGAAATATATGACGCGCAGCTGCC 2789

Qy 2764 TCGAAGACTTACCGCGAAAGCGTTATGCTGTGACGTACAAAGTCAACGAGATCCA 2823
Db 2790 TCTCAAGAGGCTGACCCCGTAAGGTGTGTATCCGTGCTGATACAGGTGAATGAATCTT 2849
Qy 2824 CTCTACGCACTTCTGACACGCTGAACGTTTACTTACACGACGAAAAAGCATT 2883
Db 2850 CTGTACGCAACCACTTCAAGAACATGTGAACGTTCTTACTGACCCGCAAGGAGCGCATC 2909
Qy 2884 GTCTGGAAGAGCTTACGTGTGTATTCCTGTGATTAAGAACATTAACGCTAAATTCGCGG 2943
Db 2910 GTGTGAAAAACATAGCGGAGCCCATGATTAATAACATGATCTGCCAAGATCCCTGG 2969
Qy 2944 GATTTACAGGCTTCAATGAGACGACTGAGGCGCAACGACGACCATTAATGAGCAGCTT 3003
Db 2970 AATTTCACTGACAGATAGAGATGTGACAGACAGTATGCTCATATGAGCAGATC 3029
Qy 3004 CTGTATTAAGCCGACAGACAGTGTATGTTCAGAAATTAAGTGAACGTCGTGCGCAAG 3063
Db 3030 TTGAGAGACCGGACCTTACCGACGTCCTTCCAGAAATTAAGCAACGTGTGTTGGCCAA 3089
Qy 3064 GCTTTAAGCCAGTCTTGGCCACGCGCAACATTTGTCTGACGAGACAGCATGTGGAGACG 3123
Db 3090 GCTTTAGTCCGAGTGTGTAAGACCGCTGGCATAGACATGACCATGTGAACATGAAACT 3149
Qy 3124 TTGACCCATTCAGACATGACAGAGCGTACTCACCTGAAATGGCACTGATCTTTTGC 3183
Db 3150 GTTGATTTATTTGAACCGACAAAGCTTCACTGACAGAGATATTAACCAACTATATC 3209
Qy 3184 ACCAGTTCCTTTGAGATGACCTTGACAGTGGGTTATTTTCGCTCTTACCGTGCACAT 3243
Db 3210 GTGAGGTTCTTTGAGACTGATCTGGAATCGCGGTCTATTTTGTGACCCAGTTCGGTTA 3265
Qy 3244 ACTTAAGGATCAGACCTGGATTAATCGCCAGGGAAGAACATGATGCGCTTAATGA 3303
Db 3270 TCCATTAGAAATATCACTGGGATTAATCCCGTGCCTTAATGATGAGGCTGAATAA 3329
Qy 3304 GAGGTAGCAAAAGGTTGTACGCGGATATCCGTGCATCACAAAAGCGGTGACACAGC 3363
Db 3330 GAAATGTCGCTCAGCTTCTGCGAGATACCAACACTGCTCGGCGAGTTGCCATGGA 3389
Qy 3364 AAGGTAGCTGATTAATGAAGATTAATACATCAAGACTACTCCAACTAATATGTGTT 3423
Db 3390 AGAGTCTATGACATGAACACTGTATACCTGGCAATTAATATCCGCGCATTAACCTATA 3449
Qy 3424 CCATTAATTCGCGGTTGCCCACTGTTGATCGTTGACCAAGACAGGGTACAACT 3483
Db 3450 CCTGTAAACAGAAAGACTGCTCATGTCTTACTCTCCACATTAATGAACACCCACAGAGT 3509
Qy 3484 GATCAGAGCGGATTCCTATCTAATGAAGAGGCAAAATCTGTGTGTGATGCGCGATCCT 3543
Db 3510 GACTTTTCTTCAATTCGTGAGCAAAATGAAGGCGCAAGATGTCTGTGTGTGGGAAAAAG 3569
Qy 3544 ATCAGCATTTCCAGGAAAGAAAGTAGAGTCATGGGTCCATTTGCCCATTAATACATCAG 3603
Db 3570 TTGTCCGTCGACAGGCAAAATGTGTGATCTGTGTGTGACAGCGGCTGAGGCTTACCT 3629
Qy 3604 TGTATCTGATTTGGGAATACCTAAGCATGTGCGTAAATATGAATTAATCTTTGTCAAT 3663
Db 3630 GCTCGGCTGATTTAAGGATCCAGGTGATGTGCCCAAAATATGACATTAATTTGTATAT 3689
Qy 3664 GTTAGAGACCCGCTACAGAAACATATCATCAACAGTGTGAGATCAGCTATCCACAC 3723
Db 3690 GTAGAGACCCATTAATTAATCAATCATATGACAGTGTGAAGACCATGCAATTAAGTT 3749
Qy 3724 AGCATGTCTAAGCTTAAGGCTGTCCACCATGAACTGACGGAACATGTGTGCTTATA 3783
Db 3750 AGCATGTGACCAAGAAAGCTTGTGTGATGTGAATCCGCGGAACCTGTGTCAAGCTA 3809
Qy 3784 GGGTGGGCTTGTGTATGCGCAACCGAAGATATCATCTGCGGTGACAGCTCATTT 3843
Db 3810 GGTATGTGTACGCTGACAGGCGCAGGAAGACATCAATGTGTGTATATGACGCGCAGTTT 3869

OY	3844	AGGTTACCCGGTCTGTCAGCTTAAGAAACCTGCCGAAAATACTAGAGTTCTCTGGTG	3903
Db	3870	AAAGTTTCCCGGATGCGAAACCCGAATCTCACTTGAAGAGACGGAAGTTCTGTTGTA	3929
OY	3904	TTCTTCGGCAAGACAAACGGCAACCAACAATGACACGAGACAGACTCGGTGTAGTGGCT	3963
Db	3930	TTCAATGGGATCGAATCGCAAGGCCCGTACGCAATCTTCAAGCTTTCAATCAACCTTG	3989
OY	3964	GACAAACATCTATCAAGGCTCAACCAAGTACGAGGACGGAAGAGCTCCAGCGTACAGAGTG	4023
Db	3990	ACCAACATTTTATACAGTTTCCAGACTCCACGAAAGCCGATGGCTGACCTCCATATCATGTG	4049
OY	4024	ATCAGAGGTGACATTAGAAGAAGCGCTACCAAGTATCTGTAATAGCTCAATTAAGCAA	4083
Db	4050	GTGCGAGGGGATTTTGCCACGCGCCACCGAAGAGATTTATTAATAGCTGTACACGCAAA	4109
OY	4084	GGTCAACACAGATTCGCGAGTGTGCGGTGCACTGTACCGAAAATGGCCGGCTCTTTGAT	4143
Db	4110	GGACAACTGGCCGAGGGGTGTGCGGACCGCTGTATAGAAATTCGCCGAAAGCTTCGAT	4165
OY	4144	AGACAGCCAAATAGCTGTGCGGACCGGTAGACTTGTGAAGCAGAACCC--GCTCATATA	4200
Db	4170	TTACAGCCGATCGAAATGAGAAAAGCGGACGTGCAAAAGGTGACGTAAACATATCATTT	4229
OY	4201	CATGCTGTAGAACCCCAATTTTTTCTTAAGATGCGCGAACCGGAGGGGACCTTAAGTCCCA	4266
Db	4230	CATGCGGTAGGACCAAACTTCAACAAAGTTTTCGAGGTTGAAGGTGAACAAACAGTTGGCA	4289
OY	4261	GCTGCTCATAGAGATAGACGTCCATCGTCAACGCTGACGGATTAACAAAATATCACTA	4320
Db	4290	GAGGTTTATGAGTCCATCGCTTAAGATTGTCAACGATATACATTAACATCAGTACGATT	4349
OY	4321	CCGCTACTGTCAACCGGATCTTATTTGTGTGGCAAAAGATCGAGTATGCAATCATTTGAT	4380
Db	4350	CCACTGTTGTCCACCGGATCTTTTCCGGGAAACAAAGATCGACTAACCCCAATCATTTGAAC	4409
OY	4381	CACCTGTTCACTGCTTTTGCACACTACGATCGCATGTCACATATATTGCTTGGATTAA	4440
Db	4410	CATTTGCTGACAGCTTTTAGACAACCATATGAGATGTAGCCATATATCTGACGGGACAAG	4465
OY	4441	CAATGGAGACCAAGATTAATCGAGGCCATTCAACCGCAAGAAAAGGCTCGAAATTTCTGGA-	4499
Db	4470	AAATGGGAAATGACTCTCAAGGAGCAAGTGGCTAGAGAGAAAGCATGGAGAGATATGC	4529
OY	4500	-----TGATGACAAAGCCAGTAGACATTGACTTGGTACAGGCTCCACCACAAAC	4545
Db	4530	ATATCCAGCACTCTTCAGTGAACAACCTGATGACAGCTGTGATGAGGTGATCCGAAG	4589
OY	4566	AGCTTTTGGCAGGACAGACCAAGTTACTTCGTCATAGAGGCAAGTTGATTTATCACTG	4605
Db	4590	AGTTCTTTGGGTGGAAAGGAGGCGCTACAGCACACAGCATGGCAAAACTTTCTCATATTTG	4649
OY	4606	GAAAGTACACATTCATCACAACCGGCCCAAGCATTTGGCGAAATCCATGCAATGTGGGCC	4665
Db	4650	GAAAGGACCAAGTTTACACGAGCGGCCCAAGGATATAGCAAAATTAATGCCATGTGGCCC	4708
OY	4666	AAACAATCTGAGGCTAATAGAGCAGATTTGCTTGTATCATCTGGGGAGAGATATGTCCAGC	4725
Db	4710	GTTGCAAGGAGGGCCAAATGAGAGGTAATGATATATCTTCGAGAAAGATGAGCAGT	4766
OY	4726	ATCGGCTCCAAATGCCAGTAGAGAGTCAAGAGCGTCTGCTCCACTCAACACTTCCA	4785
Db	4770	ATTAGGTGAAATGCCCCGTGCAAAAGTTCGAAAGCTCCACACCACTTAGACCGCTGCCT	4828
OY	4786	TGCCGTGTGTAATTAGCTATAGCGGCTACGCGGTATACAGGTTGCCGCTGTGCGAAGAA	4845
Db	4830	TGCTTGTGCATTCACATGACTGACTCCAGAAAAGTACAGCGCTTAAAGCTTCACGTCCA	4889
OY	4846	GAAACAGTTTGGCGTATGCTCATCTCTGTTGCGGAATACAGAGTACAGAGCGCTGACG	4905
Db	4890	GAAACAAATTACTGTGTGCTCATCTTTTCATTTGCCGAAATATAGAAATCACTGTGTGACG	4949
OY	4906	AAGCTACAGTGCAGCAAAACAGTCTCTGTTTTCAGCGTGTACACACGGGCTGTACACCC	4965

Db	Accession	Definition	Version	Source	Organism
Db	4950	AAAGTCCAGTCTCCAGCCGATTTGTTCTCCAGCAAGGCTCGTATATTCATCA	5009		
Oy	4966	AGGAAGTA 4973			
Db	5010	AGGAAGTA 5017			
RESULT 12					
AF004459					
LOCUS	AF004459				
DEFINITION	Venezuelan equine encephalitis virus strain 243937, complete genome.				
ACCESSION	AF004459				
VERSION	AF004459.2				
KEYWORDS	GI:5442464				
SOURCE					
ORGANISM	Venezuelan equine encephalitis virus				
REFERENCE	Venezuelan equine encephalitis virus				
AUTHORS	Vinuesa, ssRNA positive-strand viruses, no DNA stage; Togaviridae; Alphavirus; VEEV complex.				
TITLE	1 (bases 8366 to 9182)				
JOURNAL	Power, A.M., Oberste, M.S., Brault, A.C., Rico-Hesse, R.,				
MEDLINE	Schmura, S.M., Smith, J.F., Kang, W., Sweeney, W.P., and Weaver, S.C.				
PUBMED	Repeated emergence of epidemic/epizootic Venezuelan equine encephalitis from a single genotype of enzootic subtype ID virus				
REFERENCE	J. Virol. 71 (9), 6697-6705 (1997)				
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					

QY 1144 ACTAACAAATGAGAACTATCTATTAACAAGTGTGCGCCAGGCGTTTCCAGTGGGCG 1203
DB 1170 ACCAAATACCAATGAAATTAATCTTTTGGCCGTAGTGGCCAGCATTTTCTAGTGGGCA 1229
QY 1204 CGTGAACATGCTGCGCACTTGAGCAGAGAAAGAACTAGGGGTGCGGAGCGCACTTT 1263
DB 1230 AAAGAAATATAAGAAAGATGCAAGAAATGAGAGCCACTAGGAACTACAGATAGCACTTA 1289
QY 1264 ACTATGGGCTGCTGCTGGGCTTTCAAGCCAGAAATCAATCCATCTACAGAAACCT 1323
DB 1290 GTTATGGGGTGTCTGGCTTTTGAAGGCAAGATTAACATCTATTATTAAAGCGCCA 1349
QY 1324 GGTACGAAACAATTAAGAAAGTACTGCGCTTTTGACTACTATTGTGATTCAGCGCTT 1383
DB 1350 GATACCAAAACATCACTAAAGTGAACAGCATTTCCACTATTGTGTGCCAGGATA 1409
QY 1384 ACCAGCCAGGGCTGATATGAGCTTCCGCGTAGGCTCAAGCTGCTGTGAACCAACT 1443
DB 1410 GGGAGTAAACATTTGAGATCGGGCTGAACAAAGATTCAGGAAATGCTAGAAAGAGCGC 1469
QY 1444 GTCAAAACCCGACCGGCTATTACAATGCGATGAGATGAGCATCTGCGTGGCTTACAGCA 1503
DB 1470 AAGGAGCGTCACTCTCATTTATGCGAGGACGTACAGAAAGCTAAGTGGCGACCGAT 1529
QY 1504 GAAGCTGAAGAGTGGCTGCGAGCGGAAAGATCAGAGAGCCCTGCGCACTTCTCCCT 1563
DB 1530 GAGGCTTAAGAGTGGCTGAGAGCCGAGAGATTTGGCGCTGACACTACCACTTTTGGCAGCA 1589
QY 1564 GAAATAGAAAAAGAGCCGTAGAGGAGAAAGTACCTCATTTATGCAAGAGCAGAGCA 1623
DB 1590 GATGTGAGGAGCCCACTTGGAAGCCGATGTGACTTTATGCTAACAAAGAGCTGGGGCC 1649
QY 1624 GGTAGCTGAGAGAACCAACGAGACATCTAGGCTGACAACTTACCAAGGAGAAAG 1683
DB 1650 GGCTCACTGAGACACTGCGGCTTGTATTAAGGTTACAGCTATGCGGGGAGAGCAAG 1709
QY 1684 ATTGGCTTATACGCTATACCTTTCACCCAGCGGTATGAAATAGTAAAACTGGCGTGT 1743
DB 1710 ATCGGCTCTTACGAGAGCTTTTCTCCAGGCTGTACTCAAGATGAGAAATCTTTTGC 1789
QY 1744 ATCCACCATTGGCGGAAACAAGTACTGTAAATGACTCAAAAGGTAGGCGAGGAGATAC 1803
DB 1770 ATTCAACCTCTCGCTGAACAAGTCAATGATGATTAACAACCTGCGCCGAAAGGGCGTTAT 1829
QY 1804 AAGGTCAAGCCATACCAAGGTAAAGTCAATTTACCAAGAGGAGGGGGTTCCTTTCAA 1863
DB 1830 GCCGTGAACCTTACCAATGAAAAAGTGTGTGCAAGGGAGCATGCAATACCGTCCAG 1889
QY 1864 GACTTCAGGCAATGAGTGAAGAGCGCTACGATCGTTTTCACAGAGAGGAGTTGTAAC 1923
DB 1890 GACTTTCAAGCTCTGAGTGAAGATGCAACATTTGTATCAACGAGAGTTGTAAAC 1949
QY 1924 AGATACCTGCACCAATTCGAATCAACGAGAGCGCTTAAACTGACGAAGACTACTAT 1983
DB 1950 AGGATTTTGCACATATTTGCACACATGAGAGAGCGCTGAACACAGATGAAGATATATAC 2009
QY 1984 AAGACTGTAAGACTCAGAGACACAGACTCAGAAATACGCTTTCGATTTTGAACGCAAG 2043
DB 2010 AAAATTTGTCAAGCCAGAGACACAGACGCGGAATACCTGTATGACATCGACAGAAACAG 2069
QY 2044 TGTGTTAAGGAGAAAGACGAGATCCCTTGTGCTTAAACCGGTGATCTGATTCACCA 2103
DB 2070 TGCCTCAAGAAAGAACTAATGCTACCTGGGCTTAGGCTTACAGCGAGCTGTGTGATCTCC 2129
QY 2104 TTTTACGAGTTTGGTACGAGATCTCAAGACAGACGAGCAGAGACCTCAAAAGTCCCA 2163
DB 2130 TTTCAATGATTTGGCTTACGAGAGCTGAGAACAGACGCGCGCTCTTACCAAGTACCA 2189
QY 2164 ACCATCGAGTCTTGTAGAGTCCAGGTTCAAGTAAATTTGAATCATTAAGAGGCTGTG 2223
DB 2190 ACCATAGGGGTGTATGTGGGTGCGGGTTCAAGCAAGTCTGCAATCATTTAAAGCGCAGTCC 2249

QY 2224 ACTAAGAAAGATCTGTTGTGAGTGCAGAGAAAGAAATCTGCCAGAAATCATCAGGAT 2283
DB 2250 ACCAAAAAGATCTGGTGTGAGCGCCAGAAAGAAAGATCTGCCAGAAATTAAGGAGAC 2309
QY 2284 GTAAGAGATGAGACGTATGATGTGTGCTGTGAGACTGTGCATTCAGTCTTTCAAT 2343
DB 2310 GTCAAGAAATGAAGAGGCTGAGCGTCAATGCGAGAACTGTGAGCTCAGTCTTGAAT 2369
QY 2344 GGGGTTAAGCAACCCGTTAACTCTGTACATTTGATGAGGCAATTTGCTGCTCCATGAGG 2403
DB 2370 GATGCAAAACACCCCTAGAGACCTGTATATTGACGAAAGCTTTTGTCTGTACAGGC 2429
QY 2404 ACGCTGTGCACTGATTTGCAATGCTCAACCTAAGAAAGTGTATTGTGCGGAGCCCA 2463
DB 2430 ACTCTAGAGCGCTCATAGCATCATTAAGACTTAAAGGCAAGTGTCTGCGGGATCCA 2489
QY 2464 AAACATGCGGCTTTTAAATGATGTGCTGAAAGTACATTTTAAACCTATGACATATGC 2523
DB 2490 AAACAGTGGCTTTTCAATATGATGTGCTGAAAGTGCATTTTAAACCAAGATTTGC 2549
QY 2524 ACTAAGTATACATTAAGCATCTCTAGAGGAGTGCACAGACTGTAAACCGCATGCTC 2583
DB 2550 ACGAGGCTTTTCAAAAGCAATCTCTGCTGTGCACTAAATCTGTGACTTGGTGTGC 2609
QY 2584 TCCAGCTCTTCTTACGACAAAGCAATGAACCGTTAAACCATGTGCTGTAAATCATC 2643
DB 2610 TCAACCTGTTTACGACAAAGAAAGATGAAGAACAGAACCCGAAAGATTAAGATGAG 2669
QY 2644 ATAGATTCACAGGAGACACAAAGCCGACAAAGATGATTTGATTTCACTCTGTTTCA 2703
DB 2670 ATTGACACTCTGCGAGTACTTAAACCAAGCAGAGCATCTCATTTCTTTCAGA 2729
QY 2704 GATGGGTGAAGACAGTACAGATTTGACTAACAAATATCAAGAAATCATGCTGGGCTGCA 2763
DB 2730 GGGTGGTAAAGACATTTGCAATAGATTAACAAAGCAGCAAAATATAGAGGAGCTGCC 2789
QY 2764 TCCAGAGCACTTACGCGGAAAGCGTTTATGCTGTACAGTACAAAGTCAACGAAATCCA 2823
DB 2790 TCTCAAGGCTGACCGGTAAAGGCTGTATGCCGTTGGTACAAAGGATGAATAATCCC 2849
QY 2824 CTCTACTGCGACACTTGTAGCACGTGAACGTGTACTTACAGCACAGAAAAACGAT 2883
DB 2850 CTGTACGCAACCACTCAGAACATGTGAACGTCTCTGTGACCGGTACGAGGAGCCGTATC 2909
QY 2884 GTCTGAAGACGCTAGCTGTGTATCCCTGATTAAGACCTTACAGCTAAATATCCGGG 2943
DB 2910 GTGTGAAAACTAGCCGCGGATCCATGATTAATAAACTGACGCGCAATATCTCGGG 2969
QY 2944 GATTTCAAGGCTTCAATTTGAACGACTGTGACGCGGAAACGACCCATTAATGCAAGCGT 3003
DB 2970 AATTTCACTGCCAGATAGAGATGAGCAAGAGCATGAGGCAATCATGAGGACATC 3029
QY 3004 CTGTATAGCCGACAGACGTGATGTGTTCAGAAATAGGTGAACGTCTGTGGCCAG 3063
DB 3030 TTGAGAGAACCGGACCTTACGAGATTTCAAAATTAAGCGCAAGCTGTGTGGCCAG 3089
QY 3064 GCTTTAGGCAAGCTTGGCCAGCGGCAACATTTGTGTGACGAGACGAGCTGGAGAGC 3123
DB 3090 GCTTTGTGCGGATCTGAAAACTGACGAGCATGACATGACCTGAAACATGAGACACT 3149
QY 3124 TTGACCCATTTCAAGCATGACAGAGCTACTACCTGAAATGAGCACTGAATCTTTTGC 3183
DB 3150 GTGATTTATTTGAAAGAGCAAAAGCTCACTCAGCAGAGATGATGAACCAACTATGC 3209
QY 3184 ACCAGTTCTTTGAGATGACCTGAGACATGGGTTATTTTCCGCTCTACCGTGCACCT 3243
DB 3210 GTTAGTCTTTTGAAGTCACTGAGCTGAGCTCCGGCTATTTTCTGACCACTGTTCCTTA 3269
QY 3244 ACTTACAGGATCAGCATGCGGATTAACCTGCGAGGAGAAACATGTATGGGCTTAATAGA 3303
DB 3270 TCAATTTGAAATATCATCTGGATTAATTTCCCGTCCCTTACATGTAAGGCTGAATAA 3329
QY 3304 GAGGTACAAAGAGTGTGACGCGCATATCCGTGATCAAAAGCGTTGACACAGGC 3363

```

Db      3330  GAAAGGTCCTCCGACCTCCCGCAGGTACCCACAGCTGCTCGAGAGTGTCTACTGGA 3389
Qy      3364  AGGGTAGCTATATAGAAATAATACATCAAGACTACTCTCCAACTATTAATGTGTT 3423
Db      3390  AGAGCTATATGACATAAACCCGGCCACTGTGCAATATATGATCCCGCAATATCTAGTA 3449
Qy      3424  CCATTAAATCCGCGGTGCGCCACTGTGATGTTGACCAAGAAAGACAGGGTCAACT 3483
Db      3450  CCAGTAACAGAAAGCTGCTCATGTTAGTCTCCACCAATAGAAACCCACAGAGT 3509
Qy      3484  GATCAAGGGATTCCTATCTAAGATGAAGGGCAAACTGTGTGTGTATGCGGATCT 3543
Db      3510  GACTTTCTTCACTGTGTCAGCAAACTGAAGGGCAGAACTGTGTGTGTGTGCGGAGAA 3569
Qy      3544  ATCAGATTCCAGGGAAGAAAGTATAGTCCATGGGTCCATTGCCACTAATACATCAG 3603
Db      3570  TTGTCCGTCCAGGGAAGAACTGTGTATGTTGTGACACAGCTGAGGCTACCTTCAGA 3629
Qy      3604  TGTGATCTGATTTGGGAATACCTAGCCATGTCCGTAATATGACATTATCTTGTCAAT 3663
Db      3630  GCTCGCTGATTTAGTATGATCCAGTACGTGCCAATACGACATTTGATTTAT 3689
Qy      3664  GTTAGGACCCGCTACAGAAACCATCACTAACAGTGCAGATCAGCTATCCACAC 3723
Db      3690  GTGAGACTCCATATAAATATACATCATTAACAGAGTGAAGACACGCCATCAAGCTT 3749
Qy      3724  AGCATGCTAACGTGAAGGTGTCCACCACTGAAACCTGGGGAACATGTGTGCTATA 3783
Db      3750  AGCATGTGACCAAGAAAGCTGTGTGATGTGAATCCCGGGAACCTGCTCAGCAATA 3809
Qy      3784  GGGTATGGGCTTGTGATCGCGCAACCGAATATCATCTGCGGTGGCAAGCTCATTT 3843
Db      3810  GGTATGTGTACGTGACAGGGCCAGCGAAGCATTTATGTGCTATAGCGGCGAGTTC 3869
Qy      3844  AGGTTTACCCGTGTGTGACGCTTAAGAACATGCGGAAATATGAGGTTCTCTTGTG 3903
Db      3870  AAGTCTCCCGGGTATGCAACCGAATCTCATTTGAAGAGACAGAAAGTCTGTTTGA 3929
Qy      3904  TTTCTTGGCAGAGCAACGGCAACACACATGACAGAGACTGGGTGTAGTCTT 3963
Db      3930  TTCAATTGGGTACGATCGCAAGGCCCGTACGCAAACTCTTACACCTTTCATCTACTTG 3989
Qy      3964  GACAACATCTATCAAGGGTCAACAGGTACGAGGAGGAGAGCTCCAGCGTACAGAGTG 4023
Db      3990  ACCAATCTATACAGGTTTCCAGACTCCACGAAAGCCGATGGCCACTCATATATGTG 4049
Qy      4024  ATCAGAGGTGACATTAGCAAGAGCCGTGACCAAGCTATCTTAACTGTCTAATAGCAA 4083
Db      4050  GTGCGAGGGGACATTGCCAGCGCCACCGAAGAGTATCATTAATGTGCTAAGCAAGAA 4109
Qy      4084  GGTCAACCAAGTTCGGAGTGTGCGGTGACCTGTACCGAAATGGCCGGCTCTTTGAT 4143
Db      4110  GGACAACTGTGCGAGGGGTGTGCGAGCGCTGTATTAAGAAATTCGCCGAAAGCTTCAT 4169
Qy      4144  AGACAGCCCAATAGCTGTGCGGAGCGGTGACTGTGAAGACGAGAAC---GCTCATCTA 4200
Db      4170  TTACAGCCGATCGAAGTGGAAAAACGCACTGTGCAAAAGTGTACGCTAAGATATCT 4229
Qy      4201  CATGCTGATAGACCAATTTTCTAAGATGCGGAAACCGAGGGGCACTTAAAGCTCCCA 4260
Db      4230  CATGCGTAGGGCCCACTTCAACAAAGTTTCGGAAGTTGAAGTGAACAAACAGTTGGCA 4289
Qy      4261  GCTGCTTACATGAGATAGCGTCACTGTCAACGCTGAGCGGATTAACAANAATATCAGTA 4320
Db      4290  GAGGCTTATGAGTCACTGCTAATAATGTGCAACGATTAACATTAACAGTACGAGTT 4349
Qy      4321  CCGCTACTGTCAACCGGATCTATCTGTGGCAAGATCGAGTGCATCATTTGAT 4380
Db      4350  CCACGTGTGTCAACCGGATCTTCTTCGGAACAAAGATCGCTAACCAACCATTTGAAAC 4409
Qy      4381  CACCTGTCACTGCTTTCGACATACGATGCCGATGTCAACCATATATTTGCTGATAAA 4440

```

```

Db      4410  CATTTGCTGACAGCTTTTAGACACCACTGATGAGATGTAGCCATATACTGACGAGAAC 4469
Qy      4441  CAATGGAGACACGAGATATATGAGGCCATTTACCCCAAGAAAGGTGGAAATTTGAT 4500
Db      4470  AAATGGGAAATGACTCTCAAGGAAAGAGTGTCTAGAGAAAGACGTGAGATATGT 4529
Qy      4501  GATGACAAAGCC-----AGTAGCATTTGACTTGTGAGGGTCCACCAAC 4545
Db      4530  ATATCCGACAGACTCTTCGGTGAACAGAACCGGATGAGAGCTGTGATAGGTTACATCGAAG 4589
Qy      4546  AGCTTTTGGGAGGAGACAGAGTTATCTCCGTAATGAGGGCAAGTTGATTCATACCTG 4605
Db      4590  AGTTCTTGGCTGGAAGAAAGGGCTACAGACAAAGCATGCGCAACATTTCTCATATTTG 4649
Qy      4606  GAAAGTACACATTCATCAACAGCCGCAAGACATTTGCCGAAATTCATAGCATGTGGCCC 4665
Db      4650  GAAAGGACCAATTTTATCAGCGCGCCAGAGATATAGAAATTAATGATGCTATGGCCC 4709
Qy      4666  AACAAATCTGAGGCTAATGAGCAGATTGCTGTATCATCTGGGGGAGATGTGCAGC 4725
Db      4710  GTTGCAAGAGAGCCCAATGACAGATATGATGTATATCTTGTGAAAGCATGAGCAGC 4769
Qy      4726  ATTCGCTCCAATGCGCAGTAGAGAGTCAAGGCGTGTCTCCACTCACACTTCCA 4785
Db      4770  ATTAGTCAAAATGCGCGCTGGAAGAGTCCGAAGCCTCCACACCACTAGACGCTGCT 4829
Qy      4786  TGCCTGTATATAGCTATATACGCTGAGCGGCTATACAGTTCCGCTTCGGAAGAA 4845
Db      4830  TGTGTGTCATTCATGTCTATATATCCAGAAAGATACCAACGCTTAAAGCCTCAGTCCA 4889
Qy      4846  GAAAGTTCGCGCGTATGCTCATATCTGTTGGCGAAGTACAGGATCACAGGCGTGGAG 4905
Db      4890  GAACAAATTAATCTGTGTGTCTATCTTTTCATTTCCGAAAGTATAGAAATTAATCTGTG 4949
Qy      4906  AAGCTACAGTCCAGCAACCAAGTCTCTTTTTCAGCGCTGTACCAACCGGCTGTACACCCC 4965
Db      4950  AAGATCCAGTGTCTCCACCTATATCTGTTTTCACCTAAGTGTGCTGTACATTCATCA 5009
Qy      4966  AGGAAGTAC 4974
Db      5010  CGGAAGTAC 5018

RESULT 13
LOCUS      VEU55360                      11420 bp    RNA       linear    VRL 12-JUL-1999
DEFINITION Venezuelian equine encephalitis virus strain SH3, complete genome.
ACCESSION  U55360
VERSION    U55360.2  GI:5442468
KEYWORDS
SOURCE
ORGANISM  'Venezuelian equine encephalitis virus
            Venezuelian equine encephalitis virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
            Alphavirus; VEEV complex.'
REFERENCE  1 (bases 8366 to 10042)
            Weaver, S.C., Salas, R., Rico-Hesse, R., Ludwig, G.V., Oberste, M.S.,
            Boshell, J., and Tesh, R.B.
            Re-emergence of epidemic Venezuelian equine encephalomyelitis in
            South America. VEE Study Group
            Lancet 348 (9025), 436-440 (1996)
TITLE
JOURNAL   MEDLINE
PUBMED    96332169
REFERENCE  2 (bases 1 to 11420)
            Wang, E., Barrera, R., Boshell, J., Ferro, C., Freter, J.E.,
            Navarro, J.C., Salas, R., Vasquez, C., and Weaver, S.C.
            Genetic and phenotypic changes accompanying the emergence of
            enzootic subtype IC Venezuelian equine encephalitis viruses from an
            enzootic subtype ID progenitor.
            J. Virol. 73 (5), 4266-4271 (1999)
JOURNAL   MEDLINE
PUBMED    99214366
REFERENCE  3 (bases 8366 to 10042)
            Weaver, S.C.

```

TITLE Direct Submission
JOURNAL Submitted (18-APR-1996) Pathology, Univ. TX Med. Branch, 301
REFERENCE University Ave., Galveston, TX 77555-0609, USA
AUTHORS 4 (bases 1 to 11420)
Wang, B. and Weaver, S.C.
JOURNAL Direct Submission
Submitted (12-JUL-1999) Pathology, Univ. TX Med. Branch, 301
University Ave., Galveston, TX 77555-0609, USA
REMARK Sequence update by submitter
COMMENT On Jul 12, 1999 this sequence version replaced gi:3849806.
FEATURES
source
1. 11420
/organism="Venezuelan equine encephalitis virus"
/mol_type="Genomic RNA"
/strain="SH3"
/db_xref="taxon:11036"
45..7502
/note="contains nonstructural proteins 1, 2, 3 and 4;
in-frame termination codon is read through at low
frequency"
/codon_start=1
/transl_except=(pos:5658..5660,aa:OTHER)
/product="nonstructural polyprotein"
/protein_id="A043358.1"
/db_xref="GI:5442470"
/translation="MEKRVHDIEDSPFLRALORSPOFEVEAKQVTDNDHNAARF
HLSKLIETEDVPSDTLIDIGSAPARMYSKHKXHCICPMRCADPDRIYKATKLR
NCEITPEKEDKKKELAAVMSDDPLETEMTCLHDDSSREGQVAVYQVAVADPT
SLHQADKGVRAVWIGDTTPFMFKLAGAYPSYSTMADEIVLIRANGLCSDDM
ERSRKGMSILRKKYLRKSNVLFVSGSTIHEKDLRLMSVGFPLRGKQNTGTC
ETIYSCGVYVYKRIASPGIAGKSGSAATWHRGFLCKVTDLNGEYRSPCTV
PATLQDMGTGLATVDADQAKLVGNRIIVNGRTORNTMKYLLPVQAA
RMAKEYEDODERPLGRLOLVGCMARFHKITSYRPTOTLIKNSDFHSF
VLRIGSTLEIGRTIRKLEBERKPSPLTADVDQACADAKEXAREALAA
ALPRLADVEBPLEADVDFMLQAGASVEPLRGLIKIVSYADEDKISTAVISPA
VLSEKLSCHPLAEQVIVTHSGRKRAYEAPKRVHVEGALIVQDFQALSESA
TIVNEREFVNRVLIHATHGALNDEEYKIVKPEHDEYVLDIDRCKVKEELV
TGGLTGEVDPPEFAVESLTRPAAYCVPTIYGVYAGSGSKGISAVTKDL
VSAKKEACAEIIRDKMKGLDVARVTSVILNGKHPETLYIDEAPCHAGTIR
ALAIIRPKAVLCGDKQCGFPMMLCKAHFNEITQVPHKISIRCKTSYTSVS
TLFYDKMRITNPETIKIETDTGSTPKODLLITCRGVNKOLODYKNEITAA
ASQGLTRGVAVRYKVENELVATSEHNVALLTRIEDRLVMTGLADPVIKTLTA
YPGNFATIEEMQAEHAIIMHILBERPDTVFQNKANVCAKALVPLTKAGIDMT
EONTVYFETDKAHSAIYVNOICVFEGDLSGFSAATVLSIRNNHMSPP
LMVGLKEVROLSRYPRLPRAVATGVDMTGALCNDPRLINLVAPRRLPHAV
LHNHHPQSDPSSFVSKLGRTVLVGSEKLSVPKRTYDMLSDQEAIFRARDIGTG
DVERKDYVFINVRTPYKYNHYQCEDEHAIKLSMLTKACHLNGGTCVSGYAR
ASBESIIGAIAQKFSRVCKKSLBEIEVLFVFIQYDKARNTAPYKLSLTINIT
GSRLEHAGCAQSYHVGDIATGEGVIINANSKQPGGVCALYKPFESFDLP
IEVGKRLVKGAAKHIIHAVGPNFKYSEVGDROLABAYSLAKIVNDNNYKVALP
LSTGIPGKNDRILTOSLNHLTALDITDVAIYCRDKKEMTLKAVARAVEEI
CISDSSVTEPDALVYVHPKSSLAGKRYSTSGKTPSYIEGKTFQAKADIETNA
MIPVATANEQVCMYILGESMSITRSCVPESEASTPSPULPLCIHAMTPKVEQL
KASPRQITVCSFPIIKESYTRVQKIQCSPIIFSPGPVYHPRYLVITPVEEN
QSTEGTEPAVAVNDATRTMPEDIIIEEBEDSISLSDPTQVQVEADIDGPPS
VSSSGSIIPASDFDVSLSIDTLEGESYSAASATVETPFSMSPELAPAPAR
TVENPMPAPARTPTPLAPCRASSTISVTPGNAVITREELALTRBSARSA
SRTSVSNPGRVNVITREFEAFAVQOQRFDAAGTTFSSDTQGHLOQCSYQYTL
SEVULERTUEIISTAPRUDOEKELRKQDLNTPARNSYQSRVENMAITARI
LQGGHAYLKAEGVECYRTLAPVPLVSSVNRASSPSVAAECANALKEFPVLA
CIIPEYADLVMDVAGSCCLDTPAGFCPAKRSFPKISIPVLTIRSAPIQVTLN
VLAATKRNKNCVOTRELPLVDSAAFVCECKYACNNEVETKEPDIRTEENVIN
YITLKGPKAAALFAKTHNLMDODIMDRFVMDLRKDVPTGTHTEBPKQVOI
AADPLATAYICGIRRELVRLNVLNRIHTLTPMSADEPAITIAHFQBDCTLETD
IASFDKSEDDMALTLMLIEDLVADLFLILIAAFGEISHLPTKXRFAMMK
SGMFLTFVNTVINIVIASVLRRLTGSPCAPAGIDNITKVGSKLADWACRTAC
NMREYKIIDAVVEKAPYFCGFIICDTYGTACVADPLKTLFLKGLPLAADHDD
RRRLHESTRMKNVGLPELCKAVESYETVGTSSIYMAVTTLASSVKSPLYRAGP
ITLYG"
7538..11305
/note="contains capsid, E3, E2, 6K and E1 proteins"
/codon_start=1
/product="structural polyprotein"

/protein_id="AAC11999.2"
/db_xref="GI:5442469"
/translation="MFPQPMYPMQPMYRNPAPRRPFPPTDEPLAMOVBLTSS
MANULTRQRDADPEPBPAPKPKREAPQOKRGQCKKKKKCKKAKKQBPDKNS
GNKKTKTKRQKRWMLKLESDRTPIIMLEKINGACTVGGKLRPHAVEKLTND
VLAALTKRKASKYDLEADVPQWNRADPTFTEKHSIVGSHGAYQYENGRVTK
GVAKGDSGRPIIDNOGRVVAIVIGGNEBSIRALSVMMNGBGVYKYPENDEMS
LVTMCLLAVTTPPCAPPICDYRKEPETAETLAVNNGYDELLLEAAVCPKRR
STEBLEFYKLTIRPYMARICIRCAVGSCHSIIAEVYSDHDDGVLOTSGDLS
GMLKGRTRPYDMGTCIIETIRHOVLSHTSPCHIVGDFILARGCAGSITMEPK
NSYTHSCSVYEYKPNVGRLEITRPREHAEQACQYHADAQRNAVEMHLPSSV
NDSLVNVSSTVITVTPARTSALVECECGGTIKSETIRNAKQSCCTKQCRARLQ
DDKRVNNSDKLPRAAGATVILKGLHVEPLLDGKTVLADEPIITGFRVSLHLK
NPYVLTROLADPHYTHLISEPAFNVFTEKGMFVWNPKRFMAQETPAGNP
HGLPHEVITHYVADRPWSTTIGSICAIITVVAASTLFEKSRVCLTPRLTPA
RMPLCAVLCARBARFTWESLDLMNNQOMFOTLIPLAALIVTRILKVCVC
VPEPLVAGAAGAYIBHATMPSQACISINTIVNRAGVPLPISITPTKILPTYN
LEVYVCHYKTMSPALKCCGSOECTPRTYVPOCKVFTVYPPMGGACFCQTEHT
QVSKAYVMSDDCLADIAEYKATASVQAFNITVGEHSIVTVVYNGSTPVNFV
KLTAQPLSTAMPEDRKIVQAGEIYVYDEPEYAGQPGAFDIQSRTVSSDLVANT
NLVIORPKAGAIHVPTQAPSGPEOMKDKAPSLKFTAPGCEITVNPRAENCAYS
IPLAFDIPDALFTFRVSEPTPLSAECTINCCVSDPGGATVYKYSKSGKCAVHP
SGTATLKEAWEITBQSVTIHSTANIHFFRLQICTSVTCKGDCHPRDHTVTHP
QYHAQVTEAIVSSTAMTWLTSVLGGSVAIIIGLVATIVAMVTLNOKAN"

BASE COUNT 3220 a 2821 c 2895 g 2484 t
ORIGIN

Query Match 42.4%; Score 2119; DB 14; Length 11420;
Best Local Similarity 64.5%; Pred. No. 0;
Matches 3220; Conservative 0; Mismatches 1745; Indels 24; Gaps 3;

QY 10 ACTAATGATCCATATATGAAGAATACAGCTGACTTAAGTGTGACAGCCCGTATGTC 69
Db 30 AATTACTTACCCAAAATGGAGAAAGTTCAGCTTGACATCGAGAGAAGACAGCCATTCTC 89

QY 70 AAGTGTTCAGCGGACGTTTCCAAATTTGAGTGAAGCAGAGCTGACAT 129
Db 90 AGAGCTTTGCAAGGAGACTTCCGCAATTTGAGGTAGAACCAAGCAGCTACAT 149

QY 130 GACCATGCCAATGCCAAGGCTTTGGCATGTGGCAACAAAGCTCATTTGAGAGGAGTGC 189
Db 150 GACCATCTTAACGCAAGGCTTTGCGCATCTGGCTTCAAAATCGATCGAAACGAGGTG 209

QY 190 GACCGGACCAAGTATCTTGGACATTTGAAGTGCAGCTGACATGACATTTCCAT 249
Db 210 GACCATCCGACAGATCTTGAATTTGAAGTGCAGCTGACATGATTTCTTAAG 269

QY 250 CACCGTATCATTTGATCTGCTTATGATTAAGCGCTGAAGACCCGACAGACTAACAAG 309
Db 270 CATAGTATCATTTGATCTGCTGATGATGATGTCGGAAGACCCGACAGATTTGACAAG 329

QY 310 TATTCGAAGAAGCTTAAGAAAGT-----GACATTACGACAAAGACATATGCTCTAAG 363
Db 330 TATTCGAAGTAACTGAAGAAAGTAAATTCGAAGAAATTAATCTGAAGAAATGGACAGAA 389

QY 364 GCGGCAAGCTGTGTAAGTATGTCACACAGACCCAGACAGACTCCATCTGTGTATG 423
Db 390 ATGAAGAGCTTGGCGGCTGATAGAGACCTTGACCTGGAATCTGAGATGATGTGCTC 449

QY 424 CACACAGAGCCAGTGTAGTACTTTGGAAGTGTACAGTATACCAAGATGTATGCGCA 483
Db 450 CACAGACAGAGTATGCTGCTACGAAAGGCAAGTCCGCTTTTACCAGATGTATACGCA 509

QY 484 GTCCATGACAGCATCAATCTACACAGAGCGCTTAAAGAGTTAGAGCAATTTACTCG 543
Db 510 GTTACGAGACCGAAGTCTTACACCAAGCCGACAGAGGAGTTAAGTGCCTACTCG 569

QY 544 ATAGGCTTTGACAGACCCCTTTATATGACAAAACATGCGAGTTCCTACCTACTTAC 603
Db 570 ATAGGCTTTGACAGACCCCTTTATATGTTAAGAACTTGGCTGAGATATCATCTAC 629

QY 604 AACACGAAGTGGCTGACGAGAGATTTGGAAGACGTAAATGAGCTTGGCTGTAAGTCA 663

Dh 630 TCTACCAATGGGCGGACGAAAACCGTGTAAACGGCTCGTAACATAGCCCTATGACGCTCC 689
Qy 664 GATCTTCAGSAGAGAGGCTTGAAAACTCTCAATCTCTAGGAGAGAGGCTCCACCT 723
Dh 690 GACGCTACGAGCGGCTGACGTAGAGGATGTCTCATTTCTTAGAGAGAGATTGAAACCA 749
Qy 724 ACTAATAGATCATATTCGCGTTGGTTCAACATCTACAGAGAGATGATCACTGTTA 783
Dh 750 TCCAAATATGTCCTATCTCTGTGGCTGACCATCTACAGAGAGAGGACCTTACTG 809
Qy 784 CSTACTGGCATCTTCCAAACGCTGTCCCTTGAAAGAAAGTCTAATCTTACAGGTAGA 843
Dh 810 AGGACCTGGCACCTGCGCTGTATTTCACTTACGTCGCAACAAATTTACCATGTGCG 869
Qy 844 TGTGGACCAATGTGACGTGTGAGGAGTACATCAATAAAGATTAACATGACCCAGGA 903
Dh 870 TGTGAGACTATAGTTAGTTGCGACGGGTACGTGTTAAAGAAATGCTATCAGTCCAGGC 929
Qy 904 CTATACGGTAAAGTTGAGAACTTGGCGTCCACATGTCATCGGAGGGTTCTTGAGTTGC 963
Dh 930 CTGTATGGGAACCTTCAGGCTATGTGTCACGATGACCGGAGGATCTTGTGTGTC 989
Qy 964 AAAGTCAAGATACGCTGCGCGCGGAGAGGGTTTCTTTGCTGTGTATGATATGCA 1023
Dh 990 AAAGTCAAGATACGCTGCGCGCGGAGAGGGTCTTTTCCGCTGTGACGATATGCGCA 1049
Qy 1024 GCCACACTTGGCATGACATGACAGGGATCTGGCACTGAGGTAGTGTGATGACGCA 1083
Dh 1050 GCTACATTTGTGACCAAAATGCTGCGCATCTGGCAAGATGTGCTGCGGACGACGCG 1109
Qy 1084 CAAAACTATTTGGTTGGGCTCAACCAAAAGATTTGTCTCATGATGTAGACGCAAAAT 1143
Dh 1110 CAAAACTGCTGTTGGGCTCAACGAGCGCATATGCTCAACGCTGACCAAAAGAAAC 1169
Qy 1144 ACTAACAATGTCAGAACTATCTATTTACAGTGTGTCGCCAGCGCTTTTCCAGTGGCG 1203
Dh 1170 ACCAATACCATGAATAATTTCTTTTGGCCGTAGTGGCCAGGCAATTTGCTGTGGGCA 1229
Qy 1204 CGTGAACATGCTGCGCATTTGGACGACGAAAGAACTAGGGGTGCGGGAGGCACTGTT 1263
Dh 1230 AAAGAAATTAAGAAAGATCAAGAAATGAGAGGCCATGAGACTGAGATATGACATTA 1289
Qy 1264 ACTATGGGCTGCTGCTGGGCTTTCAAGACCCAGAAATCACTCATCTACAAAGAGCT 1323
Dh 1290 GTTATGGGGTGTGCTGGGCTTTTAGAAGGACAAAGATACATTAATTAAGGCCCA 1349
Qy 1324 GGTACGCAAACTAATTAAGAAAGTACCTGCGCTTTGACTCATTTTGTGATTCACGCTT 1383
Dh 1350 GATACCCAAACCATCATCAAAAGTGAACAGCGATTTCCACTATTGTGCTGCCAGATA 1409
Qy 1384 ACCAGCCACGGGCTGATATGGGCTTCCGCGTAGGCTCAAGCTGCTGTGAACCACT 1443
Dh 1410 GGCATTAACACTTGAAGATCGGGCTGAGAACAAAGATCAAGAAATGTGAAGAGCGC 1469
Qy 1444 GTCAAAACCCGACCGGCTAATTAACAATGGCGATGTGAGACTCTGCTGGCTTACAGCA 1503
Dh 1470 AAGGACCGGTACCTCTCATTTACGCGAGACGATACAAAGAGCTAAGTGGCCACCGAT 1529
Qy 1504 GAAGCTGAAGAGTGGCTGACGCGAAGAGATCAAGAAAGCCCTGCCACCTTGTCTCT 1563
Dh 1530 GAGGCTAAGGAGTGGCTGAAGCCGAGGAGTGGCGCTGCACTACCACTTTGGCAGCA 1589
Qy 1564 GAATAGAAAGAAAGACCGTGAAGGACAGAAAGTACCTCATATATGCAAGAGGACAGACA 1623
Dh 1590 GATGTTGAGAGCCCACTCTGGAAGCCAGATGTGACTTATATCAAGAGGCTGGGCGC 1649
Qy 1624 GGTACGCTGAGACACAGAGACACATCAAGGTGACAAAGTTACCCAGCAAGAGAG 1683
Dh 1650 GGCCTAGTGAAGACCTCTGTGCTTGAATTAAGGTTACAGCTATGCGCGCAGAGACAA 1709
Qy 1684 ATTGGCTCTTACGCTATACTTTCACCCAGCGGATTTGAATTAATGTAAGAAACTGCGCT 1743
Dh 1710 ATTGCTCTTACGAGTGTCTTCTCACAGGCTGTACTCAAGATGAGAAACTATCTTGC 1769

Qy 1744 ATCCACCATTTGGCGGAACAAAGTACTGGTAATGACTCAAAAGTAGGCGAGAGATAC 1803
Dh 1770 ATTCACTCTCTGCTGAACAAAGTCAATAGTATTAACACTCTGCGCCGAAAGGGCTTAT 1829
Qy 1804 AAAGTCAGCCATACCAAGTAAAGTCAATTTGACCAAGAGGACGCGGCTCTTTCAA 1863
Dh 1830 GCCGTGAACCTTACATGAAAAAGTATGTGGTCCAGAGGGACATGCAATACCCGTCAG 1889
Qy 1864 GACTTCAGGCAATTGAGTGAAGCGCTACGATCGTTTTCAGAGAGAGAGTTCGTAAC 1923
Dh 1890 GACTTCAAGCTCTGAGTGAAGGTGCAACATTTGTCAAGAAAGAGATTCGTAAC 1949
Qy 1924 AGATACCTGACCAATGCAATCAACGAGAGGCGCTAAACATGACGAAAGTACTAT 1983
Dh 1950 AGGTATTTGACCAATTTGTCACACATGAGAGGCGCTGAACACAGATGAATATTTAC 2009
Qy 1984 AAGACTGTAAAGTCAAGACACAGACTCAGAAATCGTCTTCGATATTTGACGACGAAAG 2043
Dh 2010 AAAATTTGCAAGCCAGGAGACAGCGCGAAATACGTGTATGATGACATGACGAAACAG 2069
Qy 2044 TGTGTTAAGCAAGAAAGACGAGTCCCTTGTGCTTAAACCGGTATCTGTAGATCCACA 2103
Dh 2070 TGGCTCAAGAAAGAACTAGTCACTGGGCTAGGCTTACAGGCGAGCTGTGATCTTCCC 2129
Qy 2104 TTTCAAGATTTGGCTGACAGAGTCTCAAGACAGACGAGACCTTCAAAAGTCCCA 2163
Dh 2130 TTCCATGAATTCGCTTACAGAGCTTGAAACACAGCGCGCTCTTACCAAGTACCA 2189
Qy 2164 ACCATCGAGTCTATGAGAGTGCAGGTTCAAGTAAATCTGAAATCATCAAAAGCGCTGTG 2223
Dh 2190 ACCATAGAGGTATAGGGTGTGCGGCTTCAAGCAAGTCTGGATCATTTAAAGCCAGTIC 2249
Qy 2224 ACTAAGAAAGATCTGTGTGTGATGTCGAAAGAAAGAACTGCGCAAAATCATCGGAT 2283
Dh 2250 ACCAAAAAGATCTGTGTGTGATGTCGCGCAAGAAAGAACTGCGCAAAATTAAGGAGC 2309
Qy 2284 GTAAGAGATGAGACGATGAGATGCTGCTAGAGCTGTCGATTCAGTGTCTTAAAT 2343
Dh 2310 GTCAAGAAATGAAGAGGCTGAGCTCAATGCGAATCTGTGACTCATGCTCTTGAAT 2369
Qy 2344 GGGGTTAAGCACCCGTTTAACTCTGTATCAATGATGAGGATTTGCTGCCATGACAGG 2403
Dh 2370 GGATGCAAAACCCGTATGAGACCTGTATTTGACGAAGCTTTGCTGTATGACAGG 2429
Qy 2404 ACGCTGCGCACTGATTTGCTCAACCTTAAGAAAGTGTATTTGTGCGGAGACCA 2463
Dh 2430 ACTCTCAGAGGCTCATATGCTATTAAGACCTTAAAGGAGAGTGTGCGGCGATCCA 2489
Qy 2464 AAACAAATGCGGCTTTTAAATGATGATGCGCTGAAGATCAATTTAAACATGACATATG 2523
Dh 2490 AAACAGTGGGCTTTTAAATGATGATGCTGAAGATGCAATTTTAAACACAGATTTGC 2549
Qy 2524 ACTGAAGTATCAATTAAGAGATCTCTGAGAGTGCACAGACTGTAAACGCCATCTGC 2583
Dh 2550 ACGCAGGCTTTCACAAAGAGATCTCTGTGTGTCATTAATCTGTACTTCCGCTGTC 2609
Qy 2584 TCCAGGCTTCTTACAGACAGGATGAAGACGCTTAAACCAATGTCTGTATTAATCATC 2643
Dh 2610 TCACCTGTTTTCAGCAAAAGATGAAGACGACGAAACCCGAAAGAGACTTAAGATTAG 2669
Qy 2644 ATGATACCAAGGACCAAGACCGGCAAAAGTGAATGATGATTTCACTGATTTTCGA 2703
Dh 2670 ATGACACTTACGAGTACTTAACCAAGAGAGAGATGATCTCATCTTCTGTTTCGA 2729
Qy 2704 GGAATGGTGAACAGCTACAGATTTGACATAAAATCAAGAAATCATGATGCGGCTGCA 2763
Dh 2730 GGGTGGTAAAGAGTTCGAATATGATTAACAAAGCAAGAAATTAAGACGAGCTGCC 2789
Qy 2764 TCGAAGGACTTACGCGGAAAGCGTTTATCTGTCAAGTCAAAAGTCAAGAGATCCA 2823
Dh 2790 TCTCAAGGCTGAGCCCGTAAAGGCGTGTATGCTGCTGCTCAAGGATGAAGAAATCCC 2849

OY	2824	TTCTACTGCGAGACTTCTTGAGCAGCTGTAAGCTGTTAATTCTTAACGACACAGAAAAAGCACTT	2883
Db	2850	CTGTACGACACCACCTCAGAACTGTGAACTGTCTTAACGACCCTGACGAGACCCGTATC	2909
OY	2884	GTCTGGAAGACGCTGACGTGCTGATCCCTGGATTAAGACACTTAACGCTAAATATCCGGG	2943
OY	2910	GTGTGAAAAACCTGACCGCGGATTCATGATTAATAACACTGACGGCCAAAGTACTGTGGG	2965
Db	2944	GATTTACGGCTTCAATTGACGACTGGACGGCGGAACAAGACCCATTATGCAACGCGTT	3003
Db	2970	AATTTCACTGCGACGATAGAGGAATGGCAAGACAGCAATGAGGCATCATGAGGACATC	3029
OY	3004	CTTGATTAAGCCGCGACAGACTGATGTGTTCCAGAAATAAGGTGAACGTCTGCTGGCCGAG	3063
Db	3030	TTGGAGAACCGGAGCCCTTACCGACGTATTTCAGAAATAAGGGGAACGTGTGTTGGCCCAAG	3089
OY	3064	GCTTTAGAGCCAGTCTTGAGCCACGGCGCAAACTTGCTGACGAGACACAGTGGGAGAGC	3123
Db	3090	GCTTTGGTCCGGTACTATAAACTGACAGGCACTAGACATGACCACTGAACAAATGGAACT	3149
OY	3124	TTGCACCCATTTCAAGCATGACAGAGCGTACTCACCTGAAATGCACTGMACTTCTTTGTC	3183
Db	3150	GTGATTTATTTGAAAGCAGCAAAAGCTCACTCAGACAGAGATATGATCAACAATAAGC	3209
OY	3184	ACCAGGTTCTTTGGAGTAGACTTGACAGTGGGTTATTTCCGCTCTTAACGCTGCACCT	3243
Db	3210	GTTAGGTTCTTTGGACTGCTGACCTGGACCTCCGCTCATATTTCTGCAACCCACTGTTCCGTTA	3269
OY	3244	ACTTACAGGATTCAGCACTGGGATPAACCTGCGCAGGAGAAACATGATAGGCTTAATGA	3303
Db	3270	TTCATTAGGAATTAATCACTGGGATTAATCCCGTGCCTTAACATGTACGGGCTGATTAAG	3329
OY	3304	GAGGTAGCAAGAGAGATTGTCAACGGCGATATCCGTGCATCACAAAAACGGTTGACAGAGC	3363
Db	3330	GAAGTGTCCTCGTCAAGCTCTCCCGACGATACCAACAGACTGCTCGAGCAGTGCATCTGA	3389
OY	3364	AGGGTAGCTGATTAAGGAATTAATCCATCAAGGACTCTCTCCAACTTAATATGTGTT	3423
Db	3390	AGAGCTATATGACATGMAACCGCGCACCTGTGCATTTATGATCCGCGCATTAATCTATATA	3449
OY	3424	CCATTAAATCCCGGGTTGCCCCACTCGTTGATCGTTGACCAAAAGACAGGGTACAACCT	3483
Db	3450	CCAGTGAACAGAAAGCTCTCTCATGCTTTAGTCTCCACCTAATGAACCCACAGACT	3509
OY	3484	GATCAACACGGATTCTTATCTTAAGATGAAGGCAAAATCTGTGTTGGTATCGGCGATCT	3543
Db	3510	GACTTTTTCTTATCTGACGAAACTGAAGGGCAAACTGTCTTGTTGTGTGGGGAGAG	3569
OY	3544	ATCAGCAATTCAGGGAGAAAGTAGAGTTCATGGGTCATTGGCCCACTAATACATCAAG	3603
Db	3570	TTGTCCGCTCCAGGGCAAACTGTGAATCTGTGTGACAGCAGCCGTAGAGCTTACTTCAGA	3629
OY	3604	TGTGATCTCGAATTTGGGAATACCTAGCATGTCCGTAATATGACATTAATCTTTGTCAAT	3663
Db	3630	GCTGCGCTGGATTAGTATCTCCAGGTGACGTGCCCAATACAGATTTGTAATTAAT	3689
OY	3664	GTTTAGAGCCCGGTACAGAGAACCATCACTACCAACAGTCCGAGATCAAGCTATCCACAC	3723
Db	3690	GTGAGGACTCATTAATTAATTAATCACTATTAATCAAGCTGTGAAGACACAGCATCAAGTT	3749
OY	3724	AGCATGCTAACGTGTAGAGCTGTCCACCACTGAACACTGCGCGGAACATGTGTGGTATTA	3783
Db	3750	AGCATGTTGACCAAGAAAGCTTTGTGTGATCTGAATCCCGCGGGAACCTGCGTACAGATA	3809
OY	3784	GGGATATGGGCTTGTGATCGCGGCAACCGAATAATATCACTGCGGTGGCAAGCTCATTT	3843
Db	3810	GGTATATGTTAACGCTGACAGGGCGCAAGAACATTAATTTGGTGTATAGCGCGCAAGTTC	3869
OY	3844	AGGTTTACCCGTGTGTCAAGCTTAAGAAACTGCGCAAAATTAATGAGGTTCTCTTCGTG	3903
Db	3870	AAGTTTCTCCGGGTATGCAAAACCAAAATCTCTCATCTTGAAGACAGAAAGTCTCTGTTGTA	3929
OY	3904	TTCTTCGGAAGACAAACGGCAACCAACACATGACAGGACAGACTTCGTTATGTGCTT	3963

Db	3930	TTCAATGGGTCGATCCGAAGGCCGTCACCAATCTTACAAAGCTTTTCACTTAACCTTG	3989
Qy	3964	GACAAACATCTATCAAGGGTCACACGAGTACAGGACAGGAGAGCTCCAGCGTACAGAGTG	40233
Db	3990	ACCAACATCTATACAGGTTCCAGATCCAGAGAACCGGATGCGCACCTCATTCATGCTG	4049
Qy	4024	ATCAGAGGTGACATTTAGCAGAGCGCTGACCAAGCTATCGTTAATGCTGCTATATAGCAA	40833
Db	4050	GTGCGAGGGGACATTTGCCACCGGCCACCGAAGAGATGATCAATTAATGCTGCTTACAGCAA	4109
Qy	4084	GSTCAACCAAGTTCCGGAGTGGGGTGCACTGATCCGAAAAATGGCCGGCTGCTTTGAT	41433
Db	4110	GSACAACTTGGCGAGGGGTGTGGAGACGCTGTATAGAAATATCCCGAAAGCTTGAT	4169
Qy	4144	AGACAGCCAAATAGCTGTGCGGACCGCTAGACTTGTGAAGCAGAAC---GCTCATATA	4200
Db	4170	TTACAGCCGATTCGAAGTGGGAAAAGCGGACTGATCAAAAGGTGACGATCATATCATTT	4229
Qy	4201	CATGCTGTAGAACCCCAATTTTCTTAAGATGCCGAAACCGAGGGCGACTTAACTGGCA	4260
Db	4230	CATGCCGTAGGGCCCACTTCAACAAAGTTTCGAAAGTTGAAGGTGACCAACATTTGGCA	4289
Qy	4261	GCTGCTTCAATGACGATAGCGTCCATGCTGACAGCTGACGGAGTTACAAAATATACGTA	4320
Db	4290	GAGGCTTATGAGTCCATCGCTTAAATTTGTCAACATTAACATTAAGTCACTGACGATTT	4349
Qy	4321	CCGCTACTGTCAACCCGGCATCTATTTCTGTGTGGCAAGATTCGAGTATGCAATCATTTGCAT	4380
Db	4350	CCACTGTGTCCACCGGATCTTTTCCGGGAAACAAAGATGACATTAACCAATCATTTGAC	4409
Qy	4381	CACCTGTCACTGCTTTGACACTACGATGCCATGTCACTATATTTGCTTGGATTA	4440
Db	4410	CATTTGTGACAGGCTTTTGAACCACTGATGCAATGATGACATATATCTGACGGGACAG	4469
Qy	4441	CAATGGGAGACCAAGATATATGAGGCCATTCACGGCAAGAAACGTCGAAATTTCTGGAT	4500
Db	4470	AAATGGGAAATGACTCTCAAGAGGACGTGGCTAGAGAGAGCAGTGGAGAGATATGT	4529
Qy	4501	GATGACAAAGC-----AGTAGACATTTGACTTGTCAAGGGTCCACCCAAAC	4545
Db	4530	ATATCCGACGACTCTCGGTGACAGAACCGGATGCAAGGCTGCGTAGGGTATCATCCGAG	4589
Qy	4546	AGCTCTTTGGCAGGACGACAGGTTACTCCGTCAATGAGGGCAAGTTGATTCATCCTG	4605
Db	4590	AGTCTTTGGCTGGAAGGAGGGCTACAGCAAGCATGGCAAGACTTTCCTCATATTTG	4649
Qy	4606	GAAAGTACAGATTCATACACACCGCCAAAGACATGCCAAATCCATGCAATGTGGCCC	4665
Db	4650	GAAAGGACCAAAATTTTATCAGGCGGCCCAAGGATATACAGAAATTAATGCTCATGTGCC	4709
Qy	4666	AAACAATCTGAGGCTAATGAGCAGATTTGCTTGATATCTCGGGGAGAGATATCCAGC	4725
Db	4710	GTTGCAACGGAGGCCATAGACAGTATGATATATCTTGGTGAAGATGACAGC	4769
Qy	4726	ATCCGCTTCCAATGCCCAATGAGAGAGTCAGAGGCGTCTGCTCACACTCACACTTCCA	4785
Db	4770	ATTAGGTCAAAATGCGCCGTGAGAGAGCGGAAGCTTCCACACCACTAGACCGTGGCT	4829
Qy	4786	TGCTGTGTATTTACGCTATGACGGCTGAGGGCGTATATACAGGTTGGCGCTGCGCAAGAA	4845
Db	4830	TGCTGTGTCAATCCATGCTATGATCCAGAAAGATACAGCCCTTAAAGCTTCAGTCCA	4889
Qy	4846	GAAACGTTCCGCGTATGCTCATCTTCCGTGTGCGAAGTACAGAGTACAGGCGTGACAG	4905
Db	4890	GAAACAATTAATGTTGTGCTCATCTTTCATTTGCGGAAGTATAGAAATTAATCTGGGTGCG	4949
Qy	4906	AAAGTACAGTCCAGCAAAACAGTCTGTTTTTCAAGGCGTGTATCCACCGGCTGTACACCC	4965
Db	4950	AAAGATCAGTGTCTCCAGCCTATATCTGTTCACCTTAAGGTCCTGGGTACATTCATCCA	5009
Qy	4966	AGGAATTC	4974

Db 5010 CGAAGTAC 5018

RESULT 14

LOCUS AF375051 11423 bp RNA linear VRL 26-JUN-2001

DEFINITION Venezuelan equine encephalitis virus nonstructural polypeptide and structural polypeptide genes, complete cds.

ACCESSION AF375051

VERSION AF375051.1 GI:14549692

KEYWORDS

SOURCE Venezuelan equine encephalitis virus

ORGANISM Venezuelan equine encephalitis virus

REFERENCE 1 (bases 1 to 11423)

AUTHORS Brault, A.C., Powers, A.M., Medina, G., Wang, E., Kang, W., Salas, R.A., De Siger, J., and Weaver, S.C.

TITLE Potential sources of the 1995 Venezuelan equine encephalitis subtype 1C epidemic

JOURNAL J. Virol. 75 (13), 5823-5832 (2001)

MEDLINE 21286712

PIRBASE 11390583

REFERENCE 2 (bases 1 to 11423)

AUTHORS Weaver, S.C.

TITLE Direct Submission

JOURNAL Submitted (30-Apr-2001) Pathology, Univ. Texas Medical Branch, 301 University Blvd., Galveston, TX 77555-0609, USA

FEATURES

Source

1. 11423

/organism="Venezuelan equine encephalitis virus"

/mol_type="genomic RNA"

/strain="P676"

/db_xref="taxon:11036"

/note="Isolated from mouse brain antigen"

25. 7506

/note="nsP1-4"

/codon_start=1

/transl_except="(pos:5662..5664,aa:OTHER)

/product="nonstructural polypeptide"

/protein_id="AAK66989.1"

/db_xref="GI:14549693"

/translation="MEKVVHVDIEEDSPFLALQSFPOFEVAKQVNDNDNNAFAFSLAKLITEVDSPDITLIGSAPARMAVSHKHCICIPMCAEDPDLVYATKLNCKEITDKELDKKKELAAMVSDPDLTEEMCHDDESCRGVAVQVYAADGPTSLHQANKGVAVYWGIDTTPFMKNIAGVPSYSTMADETVIAENITLCSDDVMSRRGMSILRRKYLKPSNNVLFVSGSTIYHKDLSLHMLPSVFLHGRQNTYCPATVSCDGVYVRIASISGLYKSGVATNRBGLCCYKTDITLNGERVSFVCTYVPAITCDQWGTGLATDVSDAQKLVGINORIIVNGRTORNTNWKNTLIPVQAFAFMAYEKEDODERPLGRDOLVMGCCMARFHKITISYKRPOTIITKNSDHSFVLPRIGSNTLEIGRTIRIKMLLEHKESPLITAEIDOEACADEAKEVABELRAALPLADPEEPTLEADVDLMQEGASVETPRGLIVTYSAGEDKIGSTAVLSPOAALVSEKLSCHPLAEQVIVITHSGKGVAVEPYHGVVEGNAIPVQDQALSESAITIVNERFVNRVYLHILATHGALNTDEYKTVAPSEHDEYLDIDKOCVKVELYTGIGTGLVDPPEHFAVESLRTTPAPVQVPTIGVVGSGSGSIKATVGTDLVSAKKENCAEIIIRDVKMGGLDVNARTVDSVLNGCKHPEVTLYIDAFACHTGRLALIIIRPKKAVILGSDPKQCGFNMCCIKPHNEICQVPHKISRCTSCTSVYSTLPTDKRRNTTPKTKETIEITTSSTKRDLLITCRGWKQLQIDYKNEIMTVAASQGLTRKGVAVRIKVENPLVATKSHVAVLLTRFEDRLVMTKLADPWIKTLTAKYPNFTATIEBMOAHDIAIMRIILERPPTVPOFNKAVCAKALVPLKTAGIDMTYQGNFTVDFETEDKASAEIVLNLQCFRFGDLDSGLASATVPSIINNMDNSPNNMGVANKVROLSSRRYPOLPRAVATGVTMNTGTLANYPRIINLVNRRLLHAYLHNEHPOSDFSFSKTKGRTVLVGKLSVPGKVDMLSDQEBATFRALDGIKGDVPRKDIVINRTPIYKTHIQOCEBDHAIKLSMLTKKALHLPBGTCVSIYGADRASESIIGAIARQFVRSRCKPSKSHETEVLFVGYDKRATNPYLSLTNITVYGSRLHEACAPSVHVRGDIATATEGVINANSGQGGGCVGNALKFFESPLDPIIEVGKARLVKVTKHIIVAGVNFNVSEVEDKOLAAYSIAKIVINDNYSKVAIPLLSTGISGNKDLTLOSILNHLTLALDITDAVAIYCRPKKEMTLKEAVRBAVEEICISDSSVTEPDAELVYHKSISLAGRGYSTDKTSTSYIEGTRHQAADIAEINAMPVATEANEOVCMTIIGESSSTRSKRVSESEASTPSTPLPCCLHAMPERQRLKASRPEQITVCCSPPLPKRYITGVOKIQCSOPLFSPPVPIHRRKXILTPPEVETPESEANOSTEGTEQPALVNDATIRMPERIIEEBEBSISLSLSDGPHVOYLOVEADIGSPSVSSSSWSIIPASDFVDVSLILDTLDGASVTSVASETNYSYPARMEFR

ARVPAPRTVFRNPPHAPARTPTPEPLAHSRASRTSLVSTPEGVNVRITRELEALTPSRASRSRTSLVSNPQVNRVITREFEFAVQAOQXRPAGAVIFSDPGOGLHQQKSVQTVLSREVLERTELEISYARLDOEXKELRRKQOLNPTPANSRVSREVRNKAITARRILQGLGHYILABGVCEVYRTLHPPIYLISSYNRASFSPKAVECNMLKE NPTVASICIIIPETADIVDVGASCCLDJTSFCPALKRSFKHSIYEPTIRASVPS AIQNTLQNVLAIAFCRKNQVOMELPDLTASAVNECFKAYCANNEWYFKEKPIRL LTEENVVNIYITKLGPRAALFAKTHNLNMLQDIPMRVMDLRADVETGTGTHTE RPKVNOADPLATVYLCGIRLELVNVLNLPNHTLFDMSAEDPDALIAEHFOR GDCVLETDIASFDSKEDDAMALMTILEDGVALELITLLEAAGEISLHPKTK PKRGAAMKSGMFLTPVNTVINIYIASVRERLTGSCAPAFIDDDNIYKVKSDKLM ADRCATVLMNEVKIIDAIVGKAPYFCGGFILCSVGTACRVAADPLRLKGLGPLA VDDHDDRRRLAEESTRMNRVGIILBELCAVBSRYETVGTSLIIVAMMTLASVKS FSYLRGAPITLYG"

7542. 11309

/note="encodes capsid, E2, and E1 envelope glycoproteins"

/codon_start=1

/product="structural polypeptide"

/protein_id="AAK66990.1"

/db_xref="GI:14549694"

/translation="MPPFQPMYPMQPMYRNPAPRRPMPPTDPLAMVOELTRS MANLTFKORRDAPEGPAPKPKREAPQKQGGGQKKKKQKKKATGPNPKAOS GNRKPKPKKPKRPMYKLESDFPTIMLEKINGRACVYGLKFRMHYEGKTDND VLAALKTRKASKIDLEADVPQNRADTFKTHKPOQYSMHGAVQYEGRTFVPK LGVAKGDSGRPIIDNOGRVVAIVAGVNEGSRALSVMMNKEKVTVAYTEPCNOMS LVTTMCLLANTYFPCAPPICYDRKPAETLMLSVNNDPGYDELLEAVKCPGRKR STEELPFEYKLTREPMYARCIRCAVSGSHPAIVAVNSDGHGVRLQTSOYGLDS GNLKGRMYRNMHCTEIRPLHVSILHSRPHCIYDGHGCVLLARCPGDSITMEPK GSVTSSCSVPEYKFPNVGRELTYHPREHGEQACQOYAHDAQRGAYENHLRSEV DLSLISLSSGSVTVTPVGTSLALVECKGKGISETINKAOFQSCQKECCRAARLQ NDKNVNSDKLPKAAAGATILKGLKHPVLLDAGKCTVLPAPRMIITFGRSVSLKMLRP NPTVLTTROLADEPHYTHELISEPVAENFVTEGMEFVNMHNPRAWDETLAPGN HGLPHEVITHYHRYPMSTILGISCAIVLSVAASGTWLCRSKRSCLTGYRLTPNA RMPCLAVILCCAPRABATTWESLDHLMNNQOFPVQILPLAALIVYRLLVCC VVPEPLVYAGAGAGATTHATWESQAGISNTIYNRAGYALPISTPTKILPTVN LEVYTHYKTEGMDSPAIKCCGQDCTPINRDEQCKRTGVYPRMGAGAYCFCTENT QVSKAYWMSDDCLADHAIEAYKATHASVQALNTVGBHSIVTVYVNGEPRVNGV KLTAGPSTANTPDRKIVQYAGEIIVYDFEYAGVAGDGLQSRVSSSDLYANT NLVIOBRKAGAIHVYTOAPSGFQWOMKRSKLPKTAPECEIYTNPIRBNCAVPS IPLAFDIPLALFVPSFETPLISAECTNECVSDGCIATVYKYSKSKCAVHP SGTRILKEAVALTEQSATHTHSTANIHPEFRQICTSYTCQDGHDPPODHTVHP QYHAGQTPAAASKTAMTWLTLSSGSAVIIIGLVATIVAMVYLTMQKN"

BASE COUNT 3227 a 2819 c 2902 g 2475 t

ORIGIN

Query Match 42.4%; Score 2119; DB 14; Length 11423;

Best local Similarity 64.5%; Pred. No. 0;

Matches 3223; Conservative 0; Mismatches 1750; Indels 24; Gaps 3;

2 CCTTAACTAATCGATTCGATATGGAAGAATTTCAGCTGATAGCTGACAGCC 61

2 CCCAACCATTACTCTACCCAAAATGAGAAAGTTCCAGTTCGATCGAGAAACACC 61

62 CGTATGTCAGTCTTACAGCGGAGCTTCCACAAATTTGAGATCGAAGCAAGCAGTCA 121

62 CATTCCTCAGAGCTTTACCAACGAGGCTCCCGCAGTTTGAAGTGAAGCAAGCAGTCA 121

122 CTGACATGACCAATGACCAATGACGAGGCTTTGGCATGTGGCAACAAAGCTCATTTGA 181

122 CTGTAATATGACCAATGACCAATGACGAGGCTTTGGCATGTGGCTTCAAAACTGATCGAAA 181

182 GCGAAGTGCAGCGGGGACCAAGTATTCCTTGACATTGGAAAGTGGCCCGCTGACATGAC 241

182 CGGAGGTGCAGCCATCCGACACGATCTTGACATTGGAAAGTGGCCCGCTGACATGCT 241

242 ATTCGATCAACCGCTATCATTTGATCTGCTATGATGAAGCGCTGAAGACCCGACAGAC 301

242 ATTCTAAGCATTAAGTATCATTTGATCTGCTGATGATGATGCGGAAGATCCGACAGAT 301

302 TACAACGGTATGCGAAGAAAGATTAAAGAAAGT-----GACATTAACGACAAAGAAATAG 355

302 TGTACAAGTATGACCACTAAGCTGAAGAAATAATTCGAAAGAAATAATTCGACAAAGAAATTGG 361

356 CCTCTAAGGCGGCGAGACCTGCTGGAAGTCATGTCAACCGACGACGACAGACTCATCTC 415

Db 362 ACAAGAAATGAAGAGCTGCGCCCGCTCATGAGCGACCCGAGCTGGAAATGAGACTA 421
Qy 416 TGTGTATGACACAGAGCGCAGTGTAGTACTTTGGAGTGTACAGTATACCAAGATG 475
Db 422 TGTGCTCCACGAGAGTGTATGTGCTGCTACGAGGGGCAAGTGTGTTTACCAAGATG 481
Qy 476 TGTACGAGTGTACGAGCCGACATCATCTACCAAGGGGCTTAAAGGAGTGTAGCA 535
Db 482 TATAGCGATTGAGCGACCGAAGTCTTATCACCAAGCCAAAGGAGTGTAGAGTCG 541
Qy 536 TTTATGAGTATGAGCTTTGACAGACCCCTTTTATGTACAAAGATGTGCAAGTTCCTAAC 595
Db 542 CCTATCGATTAGGCTTTGACACACCCCTTTTATGTAAAGACTGTGGCTGAGACATATC 601
Qy 596 CTACTTACACAGCACTGGGCTGACGAGAGTATTTGAAAGCAGTATACCTTGGCTCG 655
Db 602 CATCATCTCTTACCACTGGGCGCAGCAAGAACCGTGTAAACGGCTGTAACTATAGCCCTAT 661
Qy 656 GTAACTCAGATCTTACAGSAGCAGGGCTTGGAAAACCTCAATGCTTATGAGAAAGAGGC 715
Db 662 GCAGCTCCGAGCTATGAGCGGTGTCAGTAGAGGAGTCCATTTCTTAAAGAAAGTATTT 721
Qy 716 TCCAACTACTATATAGATCATATTTCTGTTGGTTCAACATCTACACAGAGATAGAT 775
Db 722 TGAACCATCCAAATATGTCTATTTCTGTGTGGTGCACATCTTACACAGAGAAAGGG 781
Qy 776 CACTGTATAGTGTGCGATTTTCCAAAGCTGTTCACCTTGAAGAAAGTCTTAACTTCA 835
Db 782 ACTTACTGAGAGGCTGGCACCTGCGCTGTATTTCACTTACGTGGCAAGCAAAATTTACA 841
Qy 836 CAGGTATATGTGGAGCACTTGTACGCTGTGAAGGATGTGATCAAAAAGTATACGATCA 895
Db 842 CATGTGCTGTGAAGCTATATGTATGTGACAGGGTACGTGTAAAGAAATAGTATCA 901
Qy 896 GCCCAGACTATATAGCTTAAAGTTGAGAACTTGGCGTCCACATGCTGCGAGGGTTCCT 955
Db 902 GTCCAGGCTGTATGTGGAGGCTTCAAGCTATATGCTGTACAGATGACCGGAGGATTCCT 961
Qy 956 TGAATTGCAAGTACAGATATGCTGCGCGCGCGAGAGGGTTCCTTGTGCTGTGTACGT 1015
Db 962 TGTCTGTCAAGGTGACAGACATTTGAAACGGGAGAGGGTCTCTTTTCCCTGTGCAAGT 1021
Qy 1016 ATGTACACAGCACTTGTGATCAGATGAGAGGATTTCTGSCAATCTGACGTTAGTGTG 1075
Db 1022 ATGTGCTCACTATTTGTGTATCAAAATGACTGGCATTCGSCAACAGATGTCAATGTGCG 1081
Qy 1076 ATGACGCAAAATCTATTTGGTGGCTCAACAAAGGATTTGCTCAATGTGTAGAGCG 1135
Db 1082 ACGACGGCAAAAACGTGTTGGGCTCAACAGCGCATATGTGTCAACGGTGTGCAACC 1141
Qy 1136 AAAAATAATCTAACATATGCAAACTATCTATTTACAGTGTGCGCCAGCGCTTTTCCA 1195
Db 1142 AAAAATACCAAAATATCTGAAGATTTATCTTTTGGCCGTAGTGGCCCAAGCATTTGTGTA 1201
Qy 1196 GTGTGGGCGCGTAATGTGTGCTGACCTTGAAGAGAAAGAACTAGGGGTGGGGAGC 1255
Db 1202 GTGTGGCAAGGAATATAGAGATTCAGAAAGTGAAGGCGCACTAGAGACTAGAGATA 1261
Qy 1256 GCACCTTACTATAGGGCTGTGCTGGGCTTTCAAGACCCAGAAATTCATCTCATCTTACA 1315
Db 1262 GACAGTTAGTATGGGGTGTGCTGGGCTTTTGAAGGCAACAAGTATCATCTATTTATA 1321
Qy 1316 AGAAGCTGTGTACGAAATTAAGAAAGTACTGCGCTCTTGTGACTCATTTGTGATTC 1375
Db 1322 AGCCCGCAGATACCCAAACCATCTCAAGTGAACAGCGATTTCCACTCATTTGCTGTCG 1381
Qy 1376 CACGCTTACAGCGCAGCGGCTGATATGGGCTTCCGCGGTAGAGTCAAGCTGCTGCTTG 1435
Db 1382 CGAGATATGGCAATGTAACACTGCGAGATCGGGCTGAGAACGGAATTCAGAAAATGCTAG 1441
Qy 1436 AACCAATGTCAAAACCCGACCGGCTATTTACATATGGCGATGTGAGCATCTGCGTGGCT 1495

Db 1442 AAGAGCAAGAGAGCGCTCACTCTCATTTACTGCGCGAGCAATATCAAGAGCTTAAGTCG 1501
Qy 1496 TACAGCAAGAAAGCTGAAGATGTGCTGCAGCGGAAAGATTCAGAGAAACCCCTGCCACCT 1555
Db 1502 CAGCGATGAGGCTTAAAGAAAGTGTGAACCCAGAGAGCTGCGCTGCATACCACTT 1561
Qy 1556 TGTCTCCCTGAATATGAAGAAAGAGACCGTGAAGGACAGATAGACCTCATTTATGCAAGAG 1615
Db 1562 TGGCAGCTGATTTTGAAGAGACCCACTGTGAACCCATGTGCACTTGTATCAAGAGG 1621
Qy 1616 CAGAGCAGGTAGCGTGGAGACACACAGAGACATCAGGGTGAACAATTTACCAGCG 1675
Db 1622 CTGGGCGCGCTCAGTGAAGACACTGTGTGCTTGAATAAAGTATACAGCTATGCCGCG 1681
Qy 1676 AAGAGAAATTTGGTCTTATGCTATTTCTTACCCGAGGGGTATTAAGTGAAGAAAC 1735
Db 1682 AGGCAAGATGTGCTTATGACAGTGTCTTCTCCACAGGCTGTATCTCAAGATGAAGAAC 1741
Qy 1736 TGGCGTATATCACCCATTTGGCGGAACAAAGTACTGTATATGACTCACAAAGGTATGGCAG 1795
Db 1742 TATTTGCAATTCACCTCTGCTGAACAACTATGTATTAACACACTCTGGCGGAAG 1801
Qy 1796 GAGATATCAAAAGTGTAGCCATATACAGGTAAAGTCAATTTGAACAAAGAGCGGCTCC 1855
Db 1802 GCGGTATGCGGTGAACCTTACCATGGAAGTATGTGTCCAGAGAGGACATGCAATAC 1861
Qy 1856 CTGTTCAGAACTTCCAGGCAATTAAGTGAAGCCCTACATCTGTTTCAAGAGAGGAGT 1915
Db 1862 CCGTCCAGACTTTCAAGCTCTGAAGTGAAGTGCACATCTGTATCAAGAACGAAAGT 1921
Qy 1916 TCGTAAACAGATACCTGACACATCGCATCAACAGAGAGGCGTAAACATGTACAGAA 1975
Db 1922 TCGTAAACAGATACCTGACACATCTGATTCACACATGAGAGAGCCCTGAACACAGATBA 1981
Qy 1976 AGTATTAAGACTGTAAAGACTTACAGACACAGCTCAAGATATCTTGTGATTTAGC 2035
Db 1982 AATATTAACAACTGTCAAGCCACAGCAGCAGCGCGAATCTGTATGACATTCAGCA 2041
Qy 2036 CAGAAAGTGTATTAACGAGAGAACGAGTCCCTTGTGCTTAAACGGTATCTGTAG 2095
Db 2042 GAAACAAATGCTCAAAAGAAATTAATCACTGGGCTGAGGCTTACAGGGAGTGTGG 2101
Qy 2096 ATCCACATTTACAGAACTTTGCGTACAGAGTCTCAAGACAGCACGACGACCTTACA 2155
Db 2102 ATCTCCCTTCCATGATTTTGCTACAGAGTGTGAACAACGTCCGGCGCTCTTACC 2161
Qy 2156 AAGTCCCAACCATCGAGCTATGAGTGTCCAGGTTCAAGTAAATCTGAATCATCAAAA 2215
Db 2162 AAGTACCAACCATGTGGGTATGTGCGTGGGCTGAGGCAAGTCTGGCATCTTAAAA 2221
Qy 2216 GCGGTGTATTAAGAAAGATCTGTTGTGAGTGTGAAGAGAAAGAACTGGGCGAAGATCA 2275
Db 2222 GCGGATTCACAAAAGATGTGTGTGAGCGCAAGAAAGAACTGGGCGAAGATTA 2281
Qy 2276 TCAAGATGTAAAGAGATGAGACGTATGATGTGTGCTAGAGCTGTGCATTCAGTGC 2335
Db 2282 TAAAGAGCTCAAGAAATATGAAGGCTGAGACGTCAATGCAAGAACTGTGACATGTC 2341
Qy 2336 TTCTAAATGSGGTTAAGCACCCCGTTAACTCTGTATCATTTGATGAGCACTTGGCTGC 2395
Db 2342 TCTTGAATGATGTCAAAACCCCGTGAAGACCTGTATTTATGAAGCAAGCTTTGCTGTC 2401
Qy 2396 ATGAGAGGACGCTGTGCACTGATTTGCCATTCGCAAACTTAAGAAAGTGTATTTGGCG 2455
Db 2402 ATGAGAGCACTTCAAGACGCTATACCATCATTAACCTTAAAGAGCAGTGTCTGCG 2461
Qy 2456 GGAAGCCAAAACAATGTGCGCTTTTAAACATGATGTCTGAAAGTACATTTTAAACATG 2515
Db 2462 GGGATCCAAAACAGTGTGCTTTTCAATATGATGTCTGAAAGTGCATTTTAAACAG 2521
Qy 2516 ACATATGCACTGAAGTATACATTAAGAACTTCTTGAAGGTGACACAGCTGTAAACG 2575
Db 2522 AGATTTCAAGAGTCTTCCACAAAGCAATCTCTGCTGTGCACTTAAATCCGTGACTT 2581

2576 GCATGCTCCACGCTCTTCTACGACAAAGCGAATGAAAGCGTTAACCCATGCTGATA 2635
2582 CGGTGCTCAACCTGTTTATCGACAAAGATGAGAACGACGAAACCGGAAAGAGACTA 2641
2656 AATATCATATGATACACAGGAGCCAAAGCCGCAAAAGATATCTGATCTTAACCT 2695
2642 AGATTGAGATTGACCTACTGCGACTACCAACCGAAGCAGACATCTCATCTTCACTT 2701
2696 GTTTCAGAGGATGGGTGAACAGCTACATGATGATCAAAAATACGAAATCATGACTG 2755
2702 GTTTCAGAGGATGGGTGAACAGCTACATGATGATGATGATGATGATGATGATGATG 2761
2756 CGGCTGATCCGCAAGGATGACCGGAAAGCGTTATGCTGCTGAGTACAAAGTCAACG 2815
2762 CAGCTGCTCCAAAGGCTGACCGGTAAGCGGTATGCGTTCGTTGATCAAGAGTGAATG 2821
2816 AGAATCCACTACTCTCGACAGCTTGTAGCAGCTGATGATGATGATGATGATGATG 2875
2822 AAAATCCCTGTACCGACCCACCTCAGAACATGATGATGATGATGATGATGATGATG 2881
2876 AACGATGCTGAGAAAGCGAGTGGATGCTGATGATGATGATGATGATGATGATGATG 2935
2882 ACCGATGCTGAGAAACACTAGCGGTGATGATGATGATGATGATGATGATGATGATG 2941
2936 ATCCGCGGATTTACCGGCTTCATTTGAGACAGCTGCGAGCGCAACAGCAGCATTTATG 2995
2942 ATCCGCGGATTTACCGGCTTCATTTGAGACAGCTGCGAGCGCAACAGCAGCATTTATG 3001
2996 CACGCTTCTTGATGATGAGCGGACAGCTGATGATGATGATGATGATGATGATGATG 3055
3002 GGCACATTTGGAGAGACCGGACCTTACCGAGTTCCTCAAAATTAAGCGAAAGTGTGT 3061
3056 GGGGGAAGGCTTTAAGCGAGCTTGGCCACCGCCAAATGCTGAGCAGACAGCAGT 3115
3062 GGGGGAAGGCTTTAAGCGAGCTTGGCCACCGCCAAATGCTGAGCAGACAGCAGT 3121
3116 GGGGGAAGGCTTTAAGCGAGCTTGGCCACCGCCAAATGCTGAGCAGACAGCAGT 3175
3122 GGAACACTGCTGATTTACTTCCAAAGGACCAAGCTCACTCAGCAGAGATGATGATGATG 3181
3176 TCTTTGCAACAGGCTTCTTGGAGTGAACCTGGAAGTGGTATTTTCCGCTCTACCG 3235
3182 AACTATGCTGAGGCTTCTTGGAGTGAACCTGGAAGTGGTATTTTCCGCTCTACCG 3241
3236 TCGACTTACTTACAGGATGACGACTGGGATTAATCCCGCAGGAAAGACATGATGGGC 3295
3242 TTCGTTATTCATTAAGAAATATCACTGGGATTAATCCCGCAGGAAAGACATGATGGGC 3301
3296 TTAATAGAGGATGACAAAGGATGTCACGGCGATTCCTGATCACTCAAAAGCGGATG 3355
3302 TGAATTAAGAGGATGTCACGGCGATTCCTGATCACTCAAAAGCGGATG 3361
3356 ACACAGGAGGATGATGATTAATTAAGAAATTAATACATCAAGGATCACTCTCAACATTA 3415
3362 CCACCGGAAAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3421
3416 ATGCTGCTTCAATTAATGCGCGGCTGCTGCTGATGATGATGATGATGATGATGATG 3475
3422 ATCTAGTACTGTAAGAAAGATGCTGATGATGATGATGATGATGATGATGATGATG 3481
3476 GTACACTGATCACAGCGGATTCCTATCTAAGATGAGGCGAAATCTGTTGGTGGATG 3535
3482 CACAGAGTATCTTTCTTCTATCTGTCAGCAACTAAGGCGAGACTGTTGGTGGATG 3541
3536 GCGATCTTATGACATTCGAGGAGAAAGTATGATGATGATGATGATGATGATGATGATG 3595
3542 GGGAGAAAGTGTCCCTCCAGGCAAAAGGTGATGATGATGATGATGATGATGATGATG 3601
3596 CCATGAGTGTGATCTGATTTGGGAAATACCTAGCAGATGCTGATTAATGATGATGATG 3655
3602 CTTTAAAGCTGCGCTGATTTAAGTATCCAGGATGATGATGATGATGATGATGATGATG 3661

3656 TTGTCATGTTAGAGACCCCGTACAGAAACATCTACTACCAACAGTGCAGATGACGCTA 3715
3662 TTATTAAGCTGAGGCTCATATATAATATCATCTATATCAGAGATGAGAACGACGCA 3721
3716 TCACACAGCATGCTTAACGTTAAAGCTGTTCACACCTGAAACCTGCGGACATGTG 3775
3722 TTAAGCTTAAGTATGTAACCAAGAAAGCTTGTGCTGATTTGAATCCCGCGAACCTCG 3781
3776 TGGCTATGAGGATGAGGCTTGTGATGCGGACACCGAATATATCACTGCGGCGGAC 3835
3782 TCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3841
3836 GCTCATTTAGGTTTAAACCGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3895
3842 GGCAGTTCAAGTTCTCCCGGATATGCAAAACCGAAATCTCAATGAAAGACAGAAATG 3901
3896 TCTTGTGTTCTTGGGCAAGGACAAAGGCAACACACATGACAGACAGATCTGCTG 3955
3902 TGTGTTGATTTGATGAGTACGATCGCAAGGCGCTGACGACATCTTACAAAGCTTTGAT 3961
3956 TAGTCTGACAAACATCTATCAAGGCTCAACAGGATGAGGCGAGGAGCTCCAGGCT 4015
3962 CTACCTTACCAACATCTATCAAGGCTCAACAGGATGAGGCGAGGAGCTCCAGGCT 4021
4016 ACAGATGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4075
4022 ATCATGTTGCTGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 4081
4076 ATGCAAAAGTCAACAGGTTCCGAGTGTGCGGTGCACTGTACCGAAATGCGCGCTG 4135
4082 ACAGCAAAAGTCAACAGGTTCCGAGTGTGCGGTGCACTGTACCGAAATGCGCGCTG 4141
4136 CTTTGTGATGAGCAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 4192
4142 GCTTGTGATGAGCAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 4201
4193 TCATCATCATGCTGTAAGGACCCAAATTTTCTTAAGATCCGGAACCGGAGGCGACTTA 4252
4202 ATATTTATCATGCTGTAAGGACCCAAATTTTCTTAAGATCCGGAACCGGAGGCGACTTA 4261
4253 AGCTGCGCTGCTTCACTGACATGAGATGAGTTCATGCTCAACGCTGAGCGGATTA 4312
4262 AGTTGCGAGAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 4321
4313 TATCAGTACCGCTACTGCTGACACCGGATTTTCTGATGAGGAAAGTGAATGATGATG 4372
4322 TAGGATTTCACTGTTGTCACCGGATTTTCTGATGAGGAAAGTGAATGATGATGATG 4381
4373 CATTCATCATGCTGTTCACTGTTGTCACCGGATTTTCTGATGAGGAAAGTGAATGATG 4432
4382 CATTCATCATGCTGTTCACTGTTGTCACCGGATTTTCTGATGAGGAAAGTGAATGATG 4441
4433 TGAATTAACATGAGGAGACAGATTAATGAGGCTTACCGCAAAAGAAAGCGTGA 4492
4442 GGGACAAAGAAATGGAATGACTCTCAAGGAGAGTGGTATGAGGAAAGAGTGAAG 4501
4493 TTTGGA-----TGATGACAAAGCAGTATGATGATGATGATGATGATGATGATG 4537
4502 AGATATGATATCAGACGACTTTCGATGAGAAACCGGATGAGGAGTGGTGAAGGATG 4561
4538 ACCCAACAGCTCTTTGGCAGGACAGACAGGTTACTCGGTCAATGAGGCAAGTGTAT 4597
4542 ATCCGAAGATTTCTTGGCTGGAAGAGGCTTACAGACAGATGATGAGGCAAGTGTAT 4621
4598 CATACCTGGAAGTATCAGATTTCACTGACCGCCCAAGGATGATGATGATGATGATG 4657
4622 CATATTTGGAAGGACCAAAATTTCACTGACCGGCGCCCAAGGATGATGATGATGATG 4681
4658 TGTGCGCCCAAAATCTGAGGCTAATGAGCAATTTGTTTATCATCTGCGGAGAGATA 4717
4682 TGTGCGCGATTTCAACGAGGCGCAATGAGCAAGTATGATGATGATGATGATGATGATG 4741
4718 TGTCCAGATCCGCTCCAAATGCGCAGATGAGGAGTCAAGAGGCTGTGCTCACTCACA 4777

OY	2	CCCTACAAACTATATGATTCGCAATATGAAAGATTCACGTTGACTTATGCTGACAGCC	61
Db	2	CCCAACCAATTAACCTTACCCAAAATGGAAGAAAGTTCACTGTATCATCGAGAAAGACAGCC	61
OY	62	CGTATGTCAAGTCGTATTACAGGGAGCGTTTCCAAATTGAGATTCGAAGCAAGGACAGTCA	121
Db	62	CATTCTCCAGAGCTTTTACACCGAGGCTTCCCGAGTTGATGAGTAAAGCAAGCAAGTCA	121
OY	122	CTGACATATGACCATGCAATGCGACAGCGTTTTCGATGTGGCAACAAAGCTCATTTGACA	181
Db	122	CTGATATATGACCATGCTATATGCGAAGCGTTTTCGATCTGGCTTCAAACTGATCGAAA	181
OY	182	GCGAAGTCGACCGGGACCAAGTTATCTTGGACATTGGAAGTCGCGCCGTCAGACATGCAC	241
Db	182	CGGAGGTGGACCCATTCGACAGATCTCTTGACATTGGAAGTCGCGCCCGCCGACAGATGT	241
OY	242	ATTCCAAATCACCGGCTATCATTTGTATCTGCCCTATGATTAAGCGCTGAAGCCCGACAGAC	301
Db	242	ATTCTAACATTAAGTATCATTTGCATCTGTCCATGAGATGTCGGAAGATTCGGACAGAT	301
OY	302	TACAACGGTATGCAGAAAGACTTAAAGAAAAGT-----GACATTTCCGACAAAGACATATG	355
Db	302	TGTACAAATATATACCTAAGTCTGAAGAAAATTTGCAAGAAAATTACTGCAAGAAATTGG	361
OY	356	CCTTAAAGCGCGACAGACTGTCTGAAGTCAATGTCAACACAGACGACAGACTCCATCTTC	415
Db	362	ACAAGAAATGAAAGAGCTCGCCGCGCTATAGAGGACCCAGACCTGGAATCTGAGACATTA	421
OY	416	TGTGTATGCACACAGACCGCACGCTGATAGTACTTTGGAAGTTAGACATATACCAAGATG	475
Db	422	TGTGCTTCCACACACATGATGTCGTGTGAAGAGGCGACATGCTGTTTTCACAGATG	481
OY	476	TGTAGCAGTCCATCAGACGACATCAATTTACACACAGCGCGTTTAAAGAGTTAGAGACA	535
Db	482	TATACGAGTTGACGAGCCGACAGACTCTTATACCAACCAACCAAGGAGTTAGATGTC	541
OY	536	TTTATCTGATAGCTTTTGAACAGACCCCTTTTATGTACAAAATAATGCGAGTGTCTTACC	595
Db	542	CCTACTGGATAGGCTTTTGAACACACCCCTTTTATGTATTAAGAACTTGGCTGAGACATATC	601
OY	596	CTACTTACAAACAGAACTGGGCTGACGAGAGTATTTGAGACACGTTAACATTTGCGCTTCG	655
Db	602	CATCATCTTACCAACTGGGCGGACGAAACCGTGTAAACGGCTCGTAAACATAGGCTCAT	661
OY	656	GTAATCTCATGATCTTACGAGAGACAGGCTTGGAAACTCATCTCTTGGAGAAAGAGGCG	715
Db	662	GCAGCTCCGACGTCATGAGGCGGTCACTGAGGAGATGTCATTTCTTTAGAGAAAGATAT	721
OY	716	TCCAACCTATATATAGATCATATTTCTGGTGTGTTCAACAACTTACACAGAAAGATAT	775
Db	722	TGAAACCATCCAAATATATGCTTATTTCTCTGTGGCTCGAACCATCTTACCGAGAAAGAGG	781
OY	776	CAGTGTACGTAGCTGGCATCTTCCAAACGTGTTCCACTTGAAGAAAGTCTTAACTTCA	835
Db	782	ACTTACTAGAGAGCTGGCACCTGCCGTGTATTTCACTTACGTGGCAAGCAAAATTACA	841
OY	836	CAGGTAGATGTGGGACCAATGTCTACGCTGTGAAGGGTACTCATCAAAAAGATTAACATCA	895
Db	842	CATGTCGGTGTGAAGCTATATGTTATTTGCGAGGGGTACGTCGTTAAAGAAATAGCTATCA	901
OY	896	GCCCAAGACTATACGCTAAAGTTTGAAGACTTGGCGTCCCAATGTGATGCGAGGGTTTCT	955
Db	902	GTTCCAGGCTGTATGGAAGCTTTCAAGGCTATGCTGTCTTACATGCAACGCGAGGGAAATCT	961
OY	956	TGAGTTGCAAAAGTCAACAATACGCTCGCGCGGAGAGGGTTTCTTTTGTGTGTACGT	1015
Db	962	TGTGTGCAAAAGTGCACAGACATTTGAACGCGGAGAGGGGTCTTTTCCGCTGTACAGT	1022
OY	1016	ATGTATCCAGCCACACTTTTGCATCAGATGACAGGGATTTTGGCACTGACGTTAGTGTGG	1075
Db	1022	ATGTGCCACGTACATTTGTGTGACCAATATGACTGGCAACAGATGTCAATGTGCGG	1081

QY	1076	ATGACGCAAAAATCTATTGGTTGGGCTGCAACCAAAAGATTGTGCTGAATGGTAGAGCGC	1135
Db	1082	ACGAGCGCCAAAACTGCTGTGGGCTCAACCAAGCCCATAGTGTAAAGGTGCAACC	1141
QY	1136	AAAGAAATATCTAAACAATGCAGAACTATATTACAGTGTGCGCCAGGCGTTTCCA	1195
Db	1142	AAAGAAACCAAAATATCCATGAAGAAATATCTTTTGCCGTAGTGGCCCAAGCATTTGCTA	1201
QY	1196	GGTGGGCGCGTGAACATGTGTCGCACTTGGACGACGAAGAAAGAACTAGAGGGGTGCGGAGC	1255
Db	1202	GGTGGGCAAAAGAAATATTAAGGAAGATCAAGAAAGATGAGAGGCCACTAGACCTACGAGATTA	1261
QY	1256	GCACTCTTACTATGGGCTGCTGCTGGGCTTTCAAGACCCGAGAAATATCATCTCACTACA	1315
Db	1262	GACAGTTATGTCATGGGGTGTGCTGGGCTTTTAAAGAGCAACAAATTAACATCTAATTTATA	1321
QY	1316	AGAAAGCTCGTAGCCAAACATTTAAGAAAGTACCTGCGCTTTGACTCATTTTGTGATTC	1375
Db	1322	AGCGCCCAAGATATCCCAACCATCATCAAAAGTGAACAGCGATTTCCATCTCATTTGCTGCTC	1381
QY	1376	CAGCGCTTACCAAGCAAGGCTCGATATGAGGCTTCCGCGCGTAGGCTCAAGCTGCTGCTTG	1435
Db	1382	CCAGGATATGCACTGATACACACTGGAAGATCGGGCTGAGAAAGAAATACGAGAAATGCTAG	1441
QY	1436	AACCAACTGTAAACCCGCAACCGGCTTTATCAATATGCCGATGTGAGACATCTGCTGGCT	1495
Db	1442	AAAGACACAAAGACCGCTCACCTTCAATTTCTGCGGAGCATTAACAAGAGCTTAAGTGC	1501
QY	1496	TACAGCAAGAGCTGAAAGAGTGGCTGACGAGGAAGATATCAGAAAGCCCTGCAACCT	1555
Db	1502	CAGCCGATGAGGCTTAAGGAAGTGGCTTAAGCCGAGAGCTGCGGCTGCATCACTACACTT	1561
QY	1556	TGCTCCCTGAAATGAAAAAGAGACCGGTAGAGGAGAAATGAGACTTATATATGACAGAG	1615
Db	1562	TGGCAGCTGATTTTGAAGAGCCCACTGTGGAAGCCGATGTGACTGATTTACAAAGAG	1621
QY	1616	CAGAGCAAGTAGGCTGAGAGACACACAGAGACATCAAGGTTGACAAATTTACCAGCGC	1675
Db	1622	CTGGGGCGGCTCAGTGGAGACACCTGTGGCTTTGATTAAGGTTTACAGCTATATGCGCGC	1681
QY	1676	AAGAAAGATTGGGCTTATGCTTACTTTTACCCGAGGCGGTATGAATAGTGAAGAAAC	1735
Db	1682	AGGACAAAGATGGCTTTACGCAAGTCTTTTCCACAGGCTGTACTCAAGAGTAGAAGAAC	1741
QY	1736	TGGGCTGATCAACCAATTTGGCGGAAAGAGTACTGTATATGACTCAAAAGTAGAGGAG	1795
Db	1742	TATCTTGATTTCAACCTCTGCTGTGAACAAAGTCAATAGTATACACACTCTGGCCGAAAG	1801
QY	1796	GAGATACAAAGTGAAGCATACCAACGCTAAGTCAATTTATACGAAGAGGACGGCGTCC	1855
Db	1802	GGCGTTATGCGGTGAAGCCCTACCATGGAAGAAAGTATGTGTCCAGAGGACATGCAATAC	1861
QY	1856	CTGTTCAGACTTTCAGAGCATTTGATGAGAGCGCTACATCTGTTTAAACGAGAGGAGT	1915
Db	1862	CCGTCAGAGATTTTCAAGCTCTGTGATGAAGTGCACATCTGTATCAAGAACGAGAGT	1921
QY	1916	TCGTAAACAGATFACCTGACCAACATTCGCAATCAACGAGAGGCGCTAAACATCGACGAG	1975
Db	1922	TCGTAAACAGATFACCTGACCAATTTTCCACACATGAGAGAGCGCTAAACACAGATGAAG	1981
QY	1976	AGTACTATAAGACTGTAAAGACTTCAGAGACACAGACTCAGATAATGCTTCGATATTGAGC	2035
Db	1982	AATATTTACAAATCTGTCAAGCCGAGGACGACGACGCGGCAATCTGTATGACTGACACA	2041
QY	2036	CACGAAGTGTATTAAAGCAGAGAGACGAGTCCCTTTGTGCTTAACCGGTGATCTGGTAG	2095
Db	2042	GGAACAATATGCTCAAAAGAAATTAATCTCATCTGGGCTATGAGGCTTACAGGAGACTGTGTG	2101
QY	2096	ATTCACCAATTTACGAGTTTGCATGACAGAGTCTCAAGACACGACGACGACCTCACA	2155
Db	2102	ATCTCTCCCTTCATGAATTTTGCTTACGAGAGTCTTGAGAACACGCTCGGCGGCTCTTTACC	2161
QY	2156	AAGTCCCAACCATGAGTCTATGAGATGTCAGAGTTCCAGTAAATCTGATATCATCAAA	2215

2162 AAGTACCAACATGAGGGGTGATGCGCGCGGGTCAGGCAAGTCGTGCAATCAATTAATA 2221
2216 GCGGTGATCTAAGAAAAGTGTGTTGTGAGTGGAAAGAGAAATCTGCGCAAGAAATCA 2275
2222 GCGGAGTCACCAAAAAGATCTGGTGTGAGCGCCAAAGAAAGAAACCTGCGCAAGAAATTA 2281
2276 TCAGGAGTGAAGAGGATGAGACGTATGATGTTGTGCTAGAGACTGTGCATTCAGTGC 2335
2282 TAAAGGACGTCAAGAAAATGAAAGGCTGAGACGTCAATGCCAAGACTGTGACCTCAAGTGC 2341
2336 TTCTAAATGGGGTTAGCAACCCCGTTAACTCTGTACATTGATGAGGCAATTTGCGCC 2395
2342 TCTTGAATGATGCAAAACACCCCGTAGAGACCTGTATATGAGCAAGCTTTTCTGTGC 2401
2396 ATGAGAGGACGCTCTGCGCACTGATTTGCCATCTGTCAAACTTAAGAAAGTGTATTGTCG 2455
2402 ATGCAAGCACTCTCAGAGCGCTCATAGCATCATAGACCTTAAAGGCAAGTGTCTGCGC 2461
2456 GGGACCCAAAACAAATGGCGCTTCTTAAATGATGCTGAAAGTACATTTAAGCAATG 2515
2462 GGGATCCAAAACAGTGTGGCTTTTCAATATGATGTGCTTAAAGTGCATTTTAAACAG 2521
2516 ACATATGCACTGAAGTGTACCATTAAGACATCTTAGAGAGTGCACACAGACTGTAAACG 2575
2522 AGATTTGACGACAGGTCTTCCAAAAGACATCTCGTGTGCACTTAATCCGTGACTT 2581
2576 CCATGCTCTCAGCGCTCTTCTACGACAAAGCAATGAAGACGGTTAAACCCATGTGCTGATA 2635
2582 CCGTGTCTCAACCTGTGTTTACACAAAGAGATGAGAACGCAACCCGAAAGACATA 2641
2636 AATATCATATGATATACCAAGAGACCAACAAAGCCGCAAAAGATGATGTGATCTTAACCT 2695
2642 AGATTTGAGATTTGACATCTGACAGTACCAACCCGAAAGAGAGAGATCTTCTACCTT 2701
2696 GTTTCAGAGAGTGGGTGAAACAGCTACGATTGACTACAAAATACAGAAATCATGACTG 2755
2702 GTTTCAGAGGGTGGGTGAAAGCAGTTGCAATATGATTAAGAAAGCAACGAAATTAATGACGG 2761
2756 CCGGTGATCGCAAGAGCTTACCGGGAAGAGCGGTTTATGTCTGATCAGGTACAAAGTCAACG 2815
2762 CAGTGTCTCTCAAGGGGTGACCGGTAAAGCGGTATGCGGTGCGTACAAAGGTGATG 2821
2816 AAGATTCACCTACTCTCGCAGACTCTGAGCAGTGAAGGTGTACTTACACGACAGAA 2875
2822 AAAATCCCTCTTACGCAACCTCAAGAACATGTGAAGCTCTTACTGACCCGCAAGAG 2881
2876 AAGCATTTGCTGGAAGACGCTAGCTGTGATCCCTGATTAAGACACTTACAGCTTAAT 2935
2882 ACCGATCGTGTGGAACCACTAGCCGGTGTATCCATGATTAATAACACTGACGGCAAGT 2941
2936 ATCCCGGGAATTTACCGGCTTCAITTTGACGACCTGGCAGCGGAACGACGCCATTTATG 2995
2942 ATCTTGGGAACCTTCACTGCAAGATGAGAAATGGCAAGACAGACATATATCCATCATGA 3001
2996 CAGCGTTCTTGATTAAGCCGAGACAGCTGATGTGTTCCAGAAATTAAGGTGAACGCTGCT 3055
3002 GGCACATCTTGAAGAGACCGAACCTTACGACGTTTTCAAAATTAAGCGCAAGCTGTGTT 3061
3056 GGGCGAAGGCTTTAGAGCCAGTCTTTGGCCACGGCCAACTGTGTGACGAGACGAGT 3115
3062 GGGCGAAGGCTTTGGTGTGCGGTACTGAAAGACTGACGAGCATAGACATGACCACTGAACAT 3121
3116 GGGGAGACGTTGACCCCATTTCAAGCATGACAGAGCTACTACCTGAATGGCACTGAACCT 3175
3122 GGAACACTGTGTGATTACTTTCAAAACGCAAAAGCTCACTGACAGAGATATGATTTGAACC 3181
3176 TCTTTTGACACAGGTTCTTTGAGATGACCTGACAGTGGGTATTTTCCGCTCTTACCG 3235
3182 AACTATGCGTAAGGTTCTTTGATCTGACCTGACTCGGCTATTTTCTGACCCACTG 3241
3336 TCGCACTTAATTAAGGATCAGACTGAGATTAATCTGCCAGGGAAGAACTGATGGGC 3295

3242 TTCCGTTATCATTAAGAAATATCACTGGATTAATTCGCCGTGCGCTAACATGTACGGGT 3301
3296 TTAATAGAGGTTAGCAAAAGAGTTGTACAGGGGATATCCGTGATCACAAAAGCGGTTG 3355
3302 TGAATTAAGAAAGTGTCCCGCACTCTCCGCAAGTACCAAACTGCTCTCGAGCACTTG 3361
3356 ACACAGGACAGGATGATGATATTAAGAAATTAATCACTCAAGGACTACTCTCAACAATTA 3415
3362 CACCGGAAGCTATGACATGAACACTGACAGCGCTGCGCAATTAATGATCCGCGCATTA 3421
3416 ATGTGTTCCATTAATGCCCGTTGCCCACTCTGTTGATCTGTGACCAAAAGACAG 3475
3422 ATCTAGACTCTGGAACAGAAAGCTGCTATCTTAAAGTCTTCCACCAATTAAGAAACCC 3481
3476 GTCAACTGATCAAGCGGATTCCTATCTAAGTGAAGGGAATATGTTGGTGTGATCG 3535
3482 CACAGAGTACTTTTCTTCAATGTGTACGAACTGAAGGGCAAGACTGTGTGGTGTGCG 3541
3536 GCGATCCATCAACATTCAGAGAAAGTGAAGTCCATGGGTCCATTGCCCATTAATA 3595
3542 GGGAGAAAGTGTCCGTCCAGGCAAAAGGTGCACTGGTTGTACAGACCAAGCTGAGGCTTA 3601
3596 CCATAGGTGTATCTGATTTGGGAATACCTTACCACTGTGCGTAAATATGACATTAATCT 3655
3602 CTTTGAAGCTCGGCTGATTTAGTATCCAGGTGACGTCGCCAAATAGACATTTGTAT 3661
3656 TTGTCAATGTTAAGAACCCGCTGACAGAAACCATCACTCAACAGTGGGAGATACGCTTA 3715
3662 TTATTAACGTAGAGACTCCATTAATTAATACATTTATCAGCAAGTGTGAAGACACAGCCA 3721
3716 TCCACACACATCTCAACGTGTAGGCTGTCCACACCTGAAACATGCGGGAACATGTG 3775
3722 TTAAGCTTATGATGTTGACCAAGAAAGCTGTGTGACTTTGAATCCGCGGGAACCTGCG 3781
3776 TGGCTATAGGATATGGCTGTGATGCGCAACCGAATATCATCTACCTCGGTGAC 3835
3782 TCAGCATAGGTATATGTTAAGTGTGACAGGCGCAGCGAGACATCATTTGCTATTAACGC 3841
3836 GCTCATTTAGGTTAACCGGTGTGCTGCTAGGCTTAAGAACATCTGCGGAATATAGAGGTTG 3895
3842 GGCAGTTCAAGTCTCCGGGATATGCAACCGAATCTTCAATGAAGAGACAGAGTAC 3901
3896 TCTTGTGTTCTTCCGCAAGAACACCGCAACACACATGACCAAGACAGACTCGGTG 3955
3902 TGTGTTATTCATTTGGGTAGATGCGAAGGCCGTGACCAATCTTACAGCTTTTCAAT 3961
3956 TAGTCTTGAACATATTATCAAGGCTCAACAGATGACAGGCAAGAGCTTCCAGCT 4015
3962 CTACCTTGAACCAATCTATATACAGGTTCCAGACTCCACGAAAGCCGATGCGCACCTCAT 4021
4016 ACAGAGTATCAGAGTGAACATTAGCAAGAGCGGTGACCAAGCTATGTTAATGTCTGTA 4075
4022 ATCATGTGTGTGCGAGGGAATTTGCCAGGCTCACGGAAGAGTATATTAATGTCTCCA 4081
4076 ATAGCAAGGTCAACAGGTTCCGAGTGTGCGGTGACGTACCGAAAAATGCGCGCTG 4135
4082 ACAGCAAGAGCAACCTGGGGAAGGCTGTGCGAGGCTGTATTAAGAAATTTCCGGAATA 4141
4136 CTTTGTATGAACAGCCAAATAGCTGTGCGAACGCTAGACTTTGAGACAGCAAGAAC--GC 4192
4142 GCTTCGATTTTACAGCCCATGGAAGTGGAAAAAGGCACTGTGTCAAAAGGTGTAACTTAAC 4201
4193 TCATCATACATAGCTGTAGGAAACCAATTTTCTAAGATGCGGGAACGGAGGGGCACTTA 4252
4202 ATATATTATCAAGCCGTAAGGCGCAACTTCAACAAAGTTTGGAAAGTTGAAGGGGCAAAAC 4261
4253 AGCTCGAGCTGCTTACATAGACATAGCTGTCAATCTGTCAACGCTGACGGAATTAACAAAA 4312
4262 AGTTGGCAAGGCTTATAGATTCATGCTAAATATGTCAAGATTAACATTAACAAAGTCAAG 4321
4313 TATCAGTACCGCTACTGTCAACGGCATCTAATTTGTGTGGCAAGATTCAGTGTGCAAT 4372
4322 TAGGATTCACACTGTTGTCCACCGGCATCTTTTCCGGGAACAAAGATGCACTAACCCAAAT 4381

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: November 15, 2003, 22:29:23 ; Search time 22769 Seconds
(without alignments)
11651.760 Million cell updates/sec

Title: US-10-023-649-1_COPY_5000_11484

Perfect score: 6485

Sequence: 1 caccgcgcgcacgcacacac.....tatttgccttaaatc 6485

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_hcg:*
- 3: gb_in:*
- 4: gb_ov:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_scs:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_scs:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_hcg_hum:*
- 31: em_hcg_inv:*
- 32: em_hcg_other:*
- 33: em_hcg_mus:*
- 34: em_hcg_pln:*
- 35: em_hcg_rtd:*
- 36: em_hcg_mam:*
- 37: em_hcg_vrt:*
- 38: em_hy:*
- 39: em_hygo_hum:*
- 40: em_hygo_mus:*
- 41: em_hygo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6485	100.0	11484	14 AF214040	AF214040 Western e
2	4075.6	62.8	4170	14 EEMSP	U03854 Western equ
3	3850.2	59.4	4047	14 AF228608	AF228608 Western e
4	2705.6	41.7	4692	14 AF339476	AF339476 Highlans
5	2397.6	37.0	11663	6 AR042411	AR042411 Sequence
6	2397.6	37.0	11663	6 AR096557	AR096557 Sequence
7	2397.6	37.0	11663	6 BD071122	BD071122 System fo
8	2396	36.9	11663	6 AR042409	AR042409 Sequence
9	2396	36.9	11663	6 AR096555	AR096555 Sequence
10	2396	36.9	11663	6 BD071120	BD071120 System fo
11	2396	36.9	11663	6 I46902	I46902 Sequence 1
12	2396	36.9	11663	14 ACU38305	U03805 Sindbis-lik
13	2388.4	36.8	11703	6 AR042412	AR042412 Sequence
14	2388.4	36.8	11703	6 AR096558	AR096558 Sequence
15	2388.4	36.8	11703	6 BD071123	BD071123 System fo
16	2386.4	36.8	11717	6 AR042410	AR042410 Sequence
17	2386.4	36.8	11717	6 AR096556	AR096556 Sequence
18	2386.4	36.8	11717	6 BD071121	BD071121 System fo
19	2386.4	36.8	11717	14 ACU38304	U03804 Sindbis-lik
20	2380.4	36.7	11703	14 SINCG	U02363 Sindbis vir
21	2380.4	36.7	11740	6 AR221250	AR221250 Sequence
22	2380.4	36.7	11740	6 AR230677	AR230677 Sequence
23	2380.4	36.7	11740	6 AR234151	AR234151 Sequence
24	2380.4	36.7	11740	6 AR237761	AR237761 Sequence
25	2380.4	36.7	11740	6 BD078845	BD078845 Recombina
26	2380.4	36.7	11740	6 BD085672	BD085672 Recombina
27	2372.2	36.6	11717	14 AF103734	AF103734 Sindbis-1
28	2369.2	36.5	16656	6 AR020969	AR020969 Sequence
29	2369.2	36.5	16656	6 AR043384	AR043384 Sequence
30	2369.2	36.5	16656	6 AR062299	AR062299 Sequence
31	2369.2	36.5	16656	6 AR183758	AR183758 Sequence
32	2365.6	36.5	11708	14 SINCK82	M09205 Ockelbo vir
33	2340.2	36.1	11569	14 AF429428	AF429428 Sindbis v
34	2329.4	35.9	11626	14 AF103728	AF103728 Sindbis v
35	2322.4	35.8	16958	6 AX65480	AX65480 Sequence
36	2227.4	34.3	11678	14 U01034	U01034 Eastern equ
37	2225.2	34.3	11687	6 AX286761	AX286761 Sequence
38	2214.2	34.1	11675	14 EEEVIRNA	X63135 Eastern Equ
39	2198.4	33.9	11684	6 AX286760	AX286760 Sequence
40	2154	33.2	4541	14 AF339474	AF339474 Buggy Cre
41	2150.8	33.2	4525	14 AF339475	AF339475 Fort Morg
42	2122.6	32.7	13905	6 AR137241	AR137241 Sequence
43	2122.6	32.7	13905	6 AR282856	AR282856 Sequence
44	1880	29.0	11824	14 AF126284	AF126284 Aura viru
45	1825.2	28.1	4676	14 AF339477	AF339477 Babanki v

ALIGNMENTS

RESULT 1
AF214040 11484 bp RNA linear VRL 27-JAN-2000
LOCUS AF214040 AF143811
DEFINITION Western equine encephalomyelitis virus strain 71V-1658
nonstructural polypeptide and structural polypeptide genes,
complete cds.
ACCESSION AF214040 AF143811
VERSION AF214040.1 GI:6760410
KEYWORDS
SOURCE
ORGANISM Western equine encephalomyelitis virus
Western equine encephalomyelitis virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
Alphavirus; WEEV complex.
REFERENCE
1 (bases 1 to 11484)
Necolitzky,D.J., Schmatz,F.L., Parker,M.D., Rayner,G.A.,

Fisher, G.R., Trent, D.W., Bader, D.E. and Nagata, L.P.
Complete genomic RNA sequence of the equine encephalomyelitis virus
and expression of the structural genes
J. Gen. Virol. 81 (Pt 1), 151-159 (2000)

JOURNAL MEDLINE PUBMED 20109043 10640553

2 (bases 1 to 11484)

Netolitzky, D.J., Schmaltz, F.L., Parker, M.D., Rayner, G.A.,
Fisher, G.R., Trent, D.W., Bader, D.E. and Nagata, L.P.
Direct Submission
Submitted (08-DEC-1999) Medical Countermeasures Section, Defence
Research Establishment Suffield, P.O. Box 4000, Stn Main, Medicine
Hat, Alberta T1A 8K6, Canada

location/Qualifiers

1. 11484

/organism="Western equine encephalomyelitis virus"

/mol_type="genomic RNA"

/strain="71V-1658"

/db_xref="taxon:11039"

1. 11484

/note="sequence of this entry added to the sequence of
AF214039, strain CBA87, constitutes the Western equine
encephalomyelitis virus complete genome"

<1. 24

25. 7428

/codon_start=1

/note="encodes nap1, nap2, nap3 and nap4 proteins"

/transl_except="(pos:5581..5583,aa:OTHER)

/product="nonstructural polyprotein"

/protein_id="AAF28339.1"

/db_xref="GI:6760411"

/translation="MERIHVLDLADSPYKSLORTPFPELEARQVTDNDNANARAF
HVAATKLESEVDROVILDIGAPVRAHNSHRHCHCPMISAEDPDRLORVARELKK
SDITDKIAKADILEVMSPTDPEATSLCHTDATCRIGSAVAVDYAVANAPSTI
YHQLKGVRTIYIGFDTPPMYKMGASPTITVNADEVLSEANIGVNSLDQSS
RLGLSLIRKKRLQPTNKIIFSVGSIYTEREGLRSMWLPNFMHLKXGKSTGRCGT
IVSCGEYVIKITTISPLGKVENLSTMEGRSLSCVMDTLLEGKSEVAVCTYVVA
TLCDOMGILATDVSVDADKLVGLNORIVNCRTORNTNOMNYLIPVAOAFSM
AREHRADLDEKEIGVERETITMGCCAFKOTKITSYKKRQITIKKVAAPFPI
PRILSHLDMGFRRLKLLLEPYKPAITMADVHLRGLQDAEEVAAAEIREAL
PPLPLEIKETIVEVDLIMOGAGSVEIPRGIHRTSYEGEKISVAIISQAVL
NSEKLAICHPLEAVLWTHKRGARVKEVHGKIVIPBGTAPVDFGLSESAVI
VFNREPVNRYLHIIAINGALNTDEEYVITKTOPDISEVPIIDARKCKEDAP
LCLTGLDVPDPHEFAVESLKTTPAAHKTPTIGVYDPSGSKGIKSVATKCDLV
SAREKNAEIIIDVRMRMDVARTVDSVILANGKRPNTVLYDEAFACHAGLIAL
IATYKPKKVVLCGDPKOCGFPMNCKLKHNRHDCSTVYHKSISRCTGYTAVLSTL
FYDRMKTNPVCADKIITDTGTYKPKHDLITCFEGWKQLOIDYKHEHMTAAS
OGLTRKGVAVRYKVENPYSQTSSEHVNLTETEKRIWKTLAGDPMWKTLYAKP
GDPTASLDWQREHDAIMARVLDKPTADVFNKVCYMAKALEPVLATANIYLRQ
WETLHPKHBAYSEMAINFCTRFAGVLDLGLFSAPVYALYRHOHNSGKYM
YGNREYAKELSRYPCTIKAVDGRVADIRNNTIKDYSPTINVPYINRRLPSLTYD
HKQGTDDHSGFLSKMGKSVLVIGDPISTPGKVESMGPIPTTITGCDLIGPSAV
GKXDIIFVNRTPYRNHYOQCEDHAIHSHMLTCKVAHNLNTHDGTCAVIGLADRYT
ENKITAVARSFRPTVCOQNTAENTEVLFVFEKDNQHNTHDGRVGLVDIYQGS
TRYEAGAPAVYVRGDISKDAQAIYVANSKQPGSGVCAKALYRKPAPDFQPIA
VGRARLYKHEPLIIVAGPNSKMBEPEDGKLAAYASIASIYNAERIKTSIVPLS
TGYSGGKDRVMOSIHLFTAPPTTDADVTYCDIKQMERIIEAHRKSEITLDD
KPDVIDIVRHNSSLAGRPYISNEKLSYLEGTFPHQTAKIATIHAMPKSEA
NEQICLYLIGESMSIRSCKPVESEASAPHTLPCNCNMAETERYIRLSAKKEP
AVCSFLPKRITIGVQLQCSKPLVSGVSPVAPVHAKAEIILPPPAATVHCN
EPTVPERIPSPVISRAPSASLISLGSVFSSASTRSSTSDVDRFVATDVHCN
TSMGISAPGLDQVLPDSVDTSHMSIPASGFEVTPSPVSDVLAECAGKRLAISE
DFNTPAPROISDYRVPVAPRRPIIPSPRSTASAPVPRPTKYQOCPGKAALEP
LDEYIRHNSXRYEAGAVITFSSGQHLQOKSROCKLOEPIIDRAVHEKYAPRLD
LEERKLOKQLOLCASEGNSRYOSRVENKATITARLLISGLTDEYVISEVNECE
VNYVPVYSSTVIRFTSAEVAVTCNVLYOENPTYASCIITDEYVADMDGAC
CLDTATCPAKLSPYKXSYLOEISAVSPIONTLOVYLAATKNCNVOMREL
PULDSAEVNDCEFKYACNDEYMDTPRDNPIRLTENTOVYVTLGKPKAALEP
NKRLOEIPMDQFMDLKDYKVTGKHKHEERKQVOVIOAADLARAAYICGIRELV
RRLNAVLLPNITLTFDMSAEDFOALIAEHFHGDPLVETDASGKSDHAIISALM
ILDEDVQDPLDLIEAAGNTSVLPTGTRFGKMGSGMELFVNTLVNIMA
SRVIRELITTSACASIGDNIIVHGVIKTLMDERCAWTNMEKXIIDAIVGKAPF
CGEFLVDDQITGTACRVADPLKRLFKLGKLPVDDTQDCRRRALHDEARNWNRIGIT

CDS

DELIVKAVESREYIILAGLIITTSITLAEYKNSIRGSPITLYG"
7473. 11183
/note="encodes nucleocapsid, E3, E2, 6K and E1 proteins"
/codon_start=1
/product="structural polyprotein"
/protein_id="AAF28340.1"
/db_xref="GI:6760412"

BASE COUNT 3251 a 2842 c 2808 g 2583 t

3' UTR

ORIGIN

Query Match 100.0%; Score 6485; DB 14; Length 11484;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6485; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CACCGCAGGACGACCAACCGTAATATGTGAACCAACTGTGCCAGAACGTATACCGCAGTC 60
5000 CACCGCAGGACGACCAACCGTAATATGTGAACCAACTGTGCCAGAACGTATACCGCAGTC 5059

61 CGGTGATTTTGAAGCAACCAAGTCCGGAATCACTGCTATGCTTGGCGGCGTCTGCTT 120

5060 CGGTGATTTTGAAGCAACCAAGTCCGGAATCACTGCTATGCTTGGCGGCGTCTGCTT 5119

121 CTAGCTTGGCAACAGCTGCTCAACCGCTCGAGGCCACTTATGACAGCGGCTTGTGGTTA 180

5120 CTAGCTTGGCAACAGCTGCTCAACCGCTCGAGGCCACTTATGACAGCGGCTTGTGGTTA 5179

181 CAGGTATGTCATCAAGCAACAGCTCTACGTGAGCATCCCTAGTGTCTCTGGCTTGG 240

5180 CAGGTATGTCATCAAGCAACAGCTCTACGTGAGCATCCCTAGTGTCTCTGGCTTGG 5239

241 ACGTCCAGCTGCTTGTGACGTACGTATTTCCACGTGAGTATTTCCAGTGCATCAGGCT 300

5240 ACGTCCAGCTGCTTGTGACGTACGTATTTCCACGTGAGTATTTCCAGTGCATCAGGCT 5299

301 TTGAAGTGAACCAACCTCTGTACAGGACCTTAACCTGGAGGTGCGCAACCTCTGGAC 360

5300 TTGAAGTGAACCAACCTCTGTACAGGACCTTAACCTGGAGGTGCGCAACCTCTGGAC 5359

361 TGGCGGAAATATGCAAGACTTAATCTGCTCTCTTCCAGTTTCTTGGACTACAGAC 420

5360 TGGCGGAAATATGCAAGACTTAATCTGCTCTCTTCCAGTTTCTTGGACTACAGAC 5419

421 CAGTACCGGACGACGAGGAGCGCCCATCCCATCAGTATGACGCGGCTTCCGACCTC 480

5420 CAGTACCGGACGACGAGGAGCGCCCATCCCATCAGTATGACGCGGCTTCCGACCTC 5479

481 CAGTTCCAAAGCCACGAGAGCTTAAGTACCAACCAACGAGAGTCCGTAAGCCATCT 540

5480 CAGTTCCAAAGCCACGAGAGCTTAAGTACCAACCAACGAGAGTCCGTAAGCCATCT 5539

Qy 541 CAGAACGGAGTTGGACGAGTACATCCGTCAACACTCCAACTGACGGTATGAAAGCGGGG 600
Db 5540 CAGAACGGAGTTGGACGAGTACATCCGTCAACACTCCAACTGACGGTATGAAAGCGGGG 5599
Qy 601 CGTATATTTTCTCATCGGAAACAGGCGCAAGTCTACCTTCAACAGAAATCAGTACGTCAAT 660
Db 5600 CGTATATTTTCTCATCGGAAACAGGCGCAAGTCTACCTTCAACAGAAATCAGTACGTCAAT 5659
Qy 661 GTAATCTCAAGAACCTATATTTGGATCGGGCCGTCATGAGAAAGTAAAGCCCGCCGCC 720
Db 5660 GTAATCTCAAGAACCTATATTTGGATCGGGCCGTCATGAGAAAGTAAAGCCCGCCGCC 5719
Qy 721 TCGATCTCGAAAGAGAGAAATGTTACAGAAAGAACTGCAATTATCGGCTCTGAGGAA 780
Db 5720 TCGATCTCGAAAGAGAGAAATGTTACAGAAAGAACTGCAATTATCGGCTCTGAGGAA 5779
Qy 781 ATGAGAGCGGTATCATATCAGAAAGTAAAGTAAAGTAAAGCAATTACAGCGAGCGAC 840
Db 5780 ATGAGAGCGGTATCATATCAGAAAGTAAAGTAAAGTAAAGCAATTACAGCGAGCGAC 5839
Qy 841 TCATTCTGGATTTGGGCAATATCTATCATCAAGAGTGAATCCTGTCGAGTGTACAGAG 900
Db 5840 TCATTCTGGATTTGGGCAATATCTATCATCAAGAGTGAATCCTGTCGAGTGTACAGAG 5899
Qy 901 TCAATTATCTGTACCAATCTACTGTCACCGTAACTTAAACAGTTTACATCTGACAGAG 960
Db 5900 TCAATTATCTGTACCAATCTACTGTCACCGTAACTTAAACAGTTTACATCTGACAGAG 5959
Qy 961 TCGCGGTTAAACGTTGCAACCTTATGTTATCAAGAAATTAACCTTACAGTACGCAATTAT 1020
Db 5960 TCGCGGTTAAACGTTGCAACCTTATGTTATCAAGAAATTAACCTTACAGTACGCAATTAT 6019
Qy 1021 GTATATACAGATGAATACAGATGCTTATCTTGAATGTTGACGCGGCTCTGCTGTACAG 1080
Db 6020 GTATATACAGATGAATACAGATGCTTATCTTGAATGTTGACGCGGCTCTGCTGTACAG 6079
Qy 1081 ATACAGCACTTTTGTCTCGGCTTAACTGAGAAAGTAAACCAAGAAAGCACTTACTTATG 1140
Db 6080 ATACAGCACTTTTGTCTCGGCTTAACTGAGAAAGTAAACCAAGAAAGCACTTACTTATG 6139
Qy 1141 AGCCAGATTAAGATCAAGCCGTCCTTATGAGAAATTAACCAATTAATTAATGATTTGG 1200
Db 6140 AGCCAGATTAAGATCAAGCCGTCCTTATGAGAAATTAACCAATTAATTAATGATTTGG 6199
Qy 1201 CTGAGCTACTTAAAGAAATGCAACGTTACCCAAATGCGAATTAATCTGCTTATGAT 1260
Db 6200 CTGAGCTACTTAAAGAAATGCAACGTTACCCAAATGCGAATTAATCTGCTTATGAT 6259
Qy 1261 CGGCGGCACTTAAATGTTGTTTCAAGAAATACGCAATGCAATGATGATCTGGGATA 1320
Db 6260 CGGCGGCACTTAAATGTTGTTTCAAGAAATACGCAATGCAATGATGATCTGGGATA 6319
Qy 1321 CCTTTGCGGATTAACCTTATTTGCGCTTACTTACAGAAAGTAAAGTAAAGTAAAGTAAAG 1380
Db 6320 CCTTTGCGGATTAACCTTATTTGCGCTTACTTACAGAAAGTAAAGTAAAGTAAAGTAAAG 6379
Qy 1381 TGAAGGGCCCAAGAACAGACAGATGTTTGGCAATCTATCTAAATCTTAAACCGTTGCAAG 1440
Db 6380 TGAAGGGCCCAAGAACAGACAGATGTTTGGCAATCTATCTAAATCTTAAACCGTTGCAAG 6439
Qy 1441 AGATACCAATGGATCAATTTGTCATGATCTTAAAGAGAGATGTCAAAGTAACTTCCCGGA 1500
Db 6440 AGATACCAATGGATCAATTTGTCATGATCTTAAAGAGAGATGTCAAAGTAACTTCCCGGA 6499
Qy 1501 CGAAACATACAGAGAGCGGCTTAAAGTGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1560
Db 6500 CGAAACATACAGAGAGCGGCTTAAAGTGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 6559
Qy 1561 CCGCTTACTTTTGGCGGATCTCATCGGGAATTAAGTCCGTGATGATGATGCGGTCTTCTG 1620
Db 6560 CCGCTTACTTTTGGCGGATCTCATCGGGAATTAAGTCCGTGATGATGATGCGGTCTTCTG 6619
Qy 1621 CAATATCATATCTCTCTTGCATGATGTCAGCGGAAGATTTTGAATGATGATGATGATGATG 1680

Db 6620 CAATATCATATCTCTCTTGCATGATGTCAGCGGAAGATTTTGAATGATGATGATGATGATG 6679
Qy 1681 ATTTCCACACGCGGACCCAGATTTGGAACGAGATCGCGTCTTGTATTAAGCGAAG 1740
Db 6680 ATTTCCACACGCGGACCCAGATTTGGAACGAGATCGCGTCTTGTATTAAGCGAAG 6739
Qy 1741 ACGAGCTATGCGCACTTTCGGGTTGATGATCTTGAAGATCTTGAAGATCTTGAAGATCTTGAAG 1800
Db 6740 ACGAGCTATGCGCACTTTCGGGTTGATGATCTTGAAGATCTTGAAGATCTTGAAGATCTTGAAG 6799
Qy 1801 TCTTAGATTTGATAGAGCGGCGGTTCCGCAATATCATCTGTGCACTTACCTTACAGGAA 1860
Db 6800 TCTTAGATTTGATAGAGCGGCGGTTCCGCAATATCATCTGTGCACTTACCTTACAGGAA 6859
Qy 1861 CGAGGTTAAATTTGGTCCATGATGAATTCGGTATGTTCTTAAAGCTTTTGTCAACA 1920
Db 6860 CGAGGTTAAATTTGGTCCATGATGAATTCGGTATGTTCTTAAAGCTTTTGTCAACA 6919
Qy 1921 CACTAGCAATATCATGATTTGCTAGCAGATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1980
Db 6920 CACTAGCAATATCATGATTTGCTAGCAGATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 6979
Qy 1981 GCGGCGCTCTATCGGCGAGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 2040
Db 6980 GCGGCGCTCTATCGGCGAGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 7039
Qy 2041 CGGAGATGCGGCACTTGGCTGAACATGAGAAATTAATTTGATGATGATGATGATGATGATGATG 2100
Db 7040 CGGAGATGCGGCACTTGGCTGAACATGAGAAATTAATTTGATGATGATGATGATGATGATGATG 7099
Qy 2101 TCAAGGACCTTACTCTGTTGGGGAATTTATCTGTTGATGATGATGATGATGATGATGATGATGAT 2160
Db 7100 TCAAGGACCTTACTCTGTTGGGGAATTTATCTGTTGATGATGATGATGATGATGATGATGATGAT 7159
Qy 2161 GCAGATGCGAGACCTCTTAAAGGCTTTTAAAGCTTGAAGAAACCAATGCGAGTCGATG 2220
Db 7160 GCAGATGCGAGACCTCTTAAAGGCTTTTAAAGCTTGAAGAAACCAATGCGAGTCGATG 7219
Qy 2221 ATACCAAGACTGCGACCGCGCGGCGGCACTGATGATGATGATGATGATGATGATGATGATGATGAT 2280
Db 7220 ATACCAAGACTGCGACCGCGCGGCGGCACTGATGATGATGATGATGATGATGATGATGATGATGAT 7279
Qy 2281 TTGAATTAAGGAGAGATTAAGAAAGGCGGTAATTCAGATGATGATGATGATGATGATGATGATGAT 2340
Db 7280 TTGAATTAAGGAGAGATTAAGAAAGGCGGTAATTCAGATGATGATGATGATGATGATGATGATGAT 7339
Qy 2341 GCTGATCATACGCTCTGTCACGTTAGCGGAAAGGCTTAAAGAACTTCAAGAGCATTA 2400
Db 7340 GCTGATCATACGCTCTGTCACGTTAGCGGAAAGGCTTAAAGAACTTCAAGAGCATTA 7399
Qy 2401 GAGGAGCCCAATCACTCTTACGCGCTGACCTTAAATAGTGAAGTGAAGTGAAGTGAAGTGAAG 2460
Db 7400 GAGGAGCCCAATCACTCTTACGCGCTGACCTTAAATAGTGAAGTGAAGTGAAGTGAAGTGAAG 7459
Qy 2461 ACCACCGGCGAGATGTTTCCATACCTTCACTTAACTTCAACAGTAACTTCAAGTAACTTCAAG 2520
Db 7460 ACCACCGGCGAGATGTTTCCATACCTTCACTTAACTTCAACAGTAACTTCAAGTAACTTCAAG 7519
Qy 2521 TCCGATGCTTACGAGATTCGAACCTCTTACGCGCGCTGAGAGCGGTTTCCGCGCC 2580
Db 7520 TCCGATGCTTACGAGATTCGAACCTCTTACGCGCGCTGAGAGCGGTTTCCGCGCC 7579
Qy 2581 GCTGCTGCTCAATTCGAAGTCTTAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2640
Db 7580 GCTGCTGCTCAATTCGAAGTCTTAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7639
Qy 2641 ATCACTTAATGCGCGCGAGGTCACCGGCAAGAAAGAAAGAAAGTGTCTTAAAGCCAA 2700
Db 7640 ATCACTTAATGCGCGCGAGGTCACCGGCAAGAAAGAAAGAAAGTGTCTTAAAGCCAA 7699
Qy 2701 AACTTACAGGCTTAAAGAAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2760

Db	7700	ACCTACTACGCTTAATAAAGAGACACGACACGCCAAGAGACGAAACCGCAAGCTTAAAC	7755
Oy	2761	AGGAAACGACACACGATGTGTATGAAGTTGGAGTCCGACACAGACATTTCCGATCATGCT	2820
Db	7760	AGGGAAACGACACACGATGTGTATGAAGTTGGAGTCCGACACAGACATTTCCGATCATGCT	7819
Oy	2821	GAACGGCCNAAGTGAATGGAATATGCTTGGCTGTCCGAGGAAGGCTGATGAAACCACTTCA	2880
Db	7820	GAACGGCCNAAGTGAATGGAATATGCTTGGCTGTCCGAGGAAGGCTGATGAAACCACTTCA	7879
Oy	2881	CGTTGAAGGAAAAAATTGATTAATGACCAATTAGCCGCGGTGAATTAAGAAAGGCTAGCAT	2940
Db	7880	CGTTGAAGGAAAAAATTGATTAATGACCAATTAGCCGCGGTGAATTAAGAAAGGCTAGCAT	7939
Oy	2941	GTACGACTTGGAGTACGGCGACGTTCCCGACAGACATGAAATCAGACAGCTGCAGTACAC	3000
Db	7940	GTACGACTTGGAGTACGGCGACGTTCCCGACAGACATGAAATCAGACAGCTGCAGTACAC	7999
Oy	3001	CAGCGACAAACCAACCGGGCTTCTACATCTGGACAACAGCGCGCATCTCCAGTATGAGAATGG	3060
Db	8000	CAGCGACAAACCAACCGGGCTTCTACATCTGGACAACAGCGCGCATCTCCAGTATGAGAATGG	8059
Oy	3061	GAGATTTACCGTACCGAGAGAGAGTGGGGGAAAGCGCACAGCGGAAACACCGATCTCTGGA	3120
Db	8060	GAGATTTACCGTACCGAGAGAGTGGGGGAAAGCGCACAGCGGAAACACCGATCTCTGGA	8119
Oy	3121	CAACAGAGGACAGAGTTGTGGCTATTGTTCTAGAGAGTCAATAGAGGACAGCGCTACCGC	3180
Db	8120	CAACAGAGGACAGAGTTGTGGCTATTGTTCTAGAGAGTCAATAGAGGACAGCGCTACCGC	8179
Oy	3181	GCTTTCAAGTGGTCACTTGGAAACAGAAAGGGGTGACATTAGGGATACCCCGGAAGTTTC	3240
Db	8180	GCTTTCAAGTGGTCACTTGGAAACAGAAAGGGGTGACATTAGGGATACCCCGGAAGTTTC	8239
Oy	3241	TGAACCGGAGTCACTAGTTTACAAGGCGTATGGGCTTTGGAATGCAAGTCCATCCATGGCA	3300
Db	8240	TGAACCGGAGTCACTAGTTTACAAGGCGTATGGGCTTTGGAATGCAAGTCCATCCATGGCA	8299
Oy	3301	CAAAACCAACCCGTGCTCTATTCACTGACGCCAGAAAGAAACACTCGACGTCCTGAAAGAA	3360
Db	8300	CAAAACCAACCCGTGCTCTATTCACTGACGCCAGAAAGAAACACTCGACGTCCTGAAAGAA	8359
Oy	3361	CGTGCACATCCAAATTACGACAGCGCTCTGAGAAAGCTTTGAAATGTCCATCACGCGC	3420
Db	8360	CGTGCACATCCAAATTACGACAGCGCTCTGAGAAAGCTTTGAAATGTCCATCACGCGC	8419
Oy	3421	GCCCAAGAAAGCAATTACCGATGACTTCAACTGACCAAGTCCCTACCTCGGGGTTCTGGCC	3480
Db	8420	GCCCAAGAAAGCAATTACCGATGACTTCAACTGACCAAGTCCCTACCTCGGGGTTCTGGCC	8479
Oy	3481	GTATTTCAGACACTCAACGCCGTGTTTACGCCAATAAATAATTGAAACGTGTGGAGCA	3540
Db	8480	GTATTTCAGACACTCAACGCCGTGTTTACGCCAATAAATAATTGAAACGTGTGGAGCA	8539
Oy	3541	ATCTGATGATGATGATGATTAAGATTCACAGTCTCGGCACAATTCCGCTACATCACGACG	3600
Db	8540	ATCTGATGATGATGATGATTAAGATTCACAGTCTCGGCACAATTCCGCTACATCACGACG	8599
Oy	3601	CACTCGCGGATGTCAACCAATTCCTGTTAATGTCTTTGACAACAGACCAATGACATCAAGGA	3660
Db	8600	CACTCGCGGATGTCAACCAATTCCTGTTAATGTCTTTGACAACAGACCAATGACATCAAGGA	8659
Oy	3661	AGACAGTATGAGAAAAATAGCTATAGCACATCTGGACCTCGCGTCTTGGCCACAA	3720
Db	8660	AGACAGTATGAGAAAAATAGCTATAGCACATCTGGACCTCGCGTCTTGGCCACAA	8719
Oy	3721	AGGGTACTTCTGTTAGCTCAATGTCTTCCAGGTGACAGTGTAAACCGTCAGTATCACGAG	3780
Db	8720	AGGGTACTTCTGTTAGCTCAATGTCTTCCAGGTGACAGTGTAAACCGTCAGTATCACGAG	8779
Oy	3781	CGAGACATCTGAAGATTATATGACACCGTGTGAGAAAAAAGATCAGAGGAAGTTGTCCGGT	3840
Db	8780	CGAGACATCTGAAGATTATATGACACCGTGTGAGAAAAAAGATCAGAGGAAGTTGTCCGGT	8839

QY	3841	GGAGAGACTGTTGTTCCACCCGTCATGGAACACTGTTAAAGTGCACCGTTTACGATCA	3900
Db	8840	AGAGAGACTGTTGTTCCACCCGTCATGGAACACTGTTAAAGTGCACCGTTTACGATCA	8899
QY	3901	CTTGAGAGAGACGTCGTCCGGGTACATTAACCATGCACAGGCCACACCGGTATTA	3968
Db	8900	CTTGAGAGAGACGTCGTCCGGGTACATTAACCATGCACAGGCCACACCGGTATTA	8959
QY	3961	GTCTATTTGGAGGAAAGCTGAGCGCAAGTATACATTAAACACTTCTGGCAAGAACGT	4020
Db	8960	GTCTATTTGGAGGAAAGCTGAGCGCAAGTATACATTAAACACTTCTGGCAAGAACGT	9019
QY	4021	CACCTAGCAATGTAAGTGTGCGACCTACAGACAGAGATCGTGAAGCAGCAAGAAAT	4080
Db	9020	CACCTAGCAATGTAAGTGTGCGACCTACAGACAGAGATCGTGAAGCAGCAAGAAAT	9079
QY	4081	GAACGGCTGCACCTAAGCAAAACAGTGCAATTCCTTACAGAGCAACCAAGAAATGGT	4140
Db	9080	GAACGGCTGCACCTAAGCAAAACAGTGCAATTCCTTACAGAGCAACCAAGAAATGGT	9139
QY	4141	CTTCAACTCCGCGGATCTTATTAGGACACAGACACACTGAGTGAAGTAAATTGCACAT	4200
Db	9140	CTTCAACTCCGCGGATCTTATTAGGACACAGACACACTGAGTGAAGTAAATTGCACAT	9199
QY	4201	TCCATTCCGCTTGCACACGCACAGTCTGCGCGGTTCCGTTAGCTCACAACGCTTACAGTCAC	4260
Db	9200	TCCATTCCGCTTGCACACGCACAGTCTGCGCGGTTCCGTTAGCTCACAACGCTTACAGTCAC	9259
QY	4261	GAAGTGTGTTAAAGCATCAACCTTCCACTGACCTGCAATGGACCAACTTGTGCAGAAC	4320
Db	9260	GAAGTGTGTTAAAGCATCAACCTTCCACTGACCTGCAATGGACCAACTTGTGCAGAAC	9319
QY	4321	GAGAAAAATTGGGGCTGCGAGACAGACGAACAGCAATGATTAACAGGTTCTACATCCAG	4380
Db	9320	GAGAAAAATTGGGGCTGCGAGACAGACGAACAGCAATGATTAACAGGTTCTACATCCAG	9379
QY	4381	GAATTTTTCTGTGGGGCGAGAAAGGCTGAGTACGTAATGGGTTAACATGAACCAAGTCAG	4440
Db	9380	GAATTTTTCTGTGGGGCGAGAAAGGCTGAGTACGTAATGGGTTAACATGAACCAAGTCAG	9439
QY	4441	AGTCGGGGCCCAAGAGTCCGACCAAGGCGAACCAATGAGTGGCGATGATCATCAT	4500
Db	9440	AGTCGGGGCCCAAGAGTCCGACCAAGGCGAACCAATGAGTGGCGATGATCATCAT	9499
QY	4501	CCACTATTATCATCGGCATCCAGTCTTCACTGTCAATGTGTGTGTGTGTGTGTGTGTGTGC	4560
Db	9500	CCACTATTATCATCGGCATCCAGTCTTCACTGTCAATGTGTGTGTGTGTGTGTGTGTGTGC	9559
QY	4561	TATCTCTGTTAGGCACTGCATATATAGACAGTGTGATGGCCAAAGCAAGAAAGACTGCTCT	4620
Db	9560	TATCTCTGTTAGGCACTGCATATATAGACAGTGTGATGGCCAAAGCAAGAAAGACTGCTCT	9619
QY	4621	GACCCCATACCGCGTTTGCACCGAAGCAAGCGTATCCCAAGCATTAACGCGTTTGTGCTG	4680
Db	9620	GACCCCATACCGCGTTTGCACCGAAGCAAGCGTATCCCAAGCATTAACGCGTTTGTGCTG	9679
QY	4681	CATTGCGCAACCAACGCTGAACATTGGAGAAACTTTGAACCATCTGTGCTTTAACAA	4740
Db	9680	CATTGCGCAACCAACGCTGAACATTGGAGAAACTTTGAACCATCTGTGCTTTAACAA	9739
QY	4741	CCAACCGTTTCTCGGGGACAGTTGTGCAATTCCTGTGACAGCGTTGTATTCGTTCGG	4800
Db	9740	CCAACCGTTTCTCGGGGACAGTTGTGCAATTCCTGTGACAGCGTTGTATTCGTTCGG	9799
QY	4801	CTGCTTTTCAATGCTGCAATGCTTTTTTATTGTTGACAGGCGTCTGCTGGGGAAGTAGA	4860
Db	9800	CTGCTTTTCAATGCTGCAATGCTTTTTTATTGTTGACAGGCGTCTGCTGGGGAAGTAGA	9859
QY	4861	CGCCTTTGAACATGCGACCACTGTGCGCAAAATGTTCCGGGGATCCCGTTAAGAGCGTTGCT	4920
Db	9860	CGCCTTTGAACATGCGACCACTGTGCGCAAAATGTTCCGGGGATCCCGTTAAGAGCGTTGCT	9919

QY 4921 CGAACGCGAGTTACGCGCCACCTTAACCTGAGATCAAGGTCGTCTCATCGGAATTAC 4980
Db 9920 CGAACGCGAGTTACGCGCCACCTTAACCTGAGATCAAGGTCGTCTCATCGGAATTAC 9979
QY 4981 ACCTTCACTAACAAGAGTACGTACCTGCAAAATTTCCACAGATCATTTCTTCCACCA 5040
Db 9980 ACCTTCACTAACAAGAGTACGTACCTGCAAAATTTCCACAGATCATTTCTTCCACCA 10039
QY 5041 AGTTAAATGTCGCGGTCCTCTGAGTGCAGGATCTCAAGGCGGATTAACATGCGG 5100
Db 10040 AGTTAAATGTCGCGGTCCTCTGAGTGCAGGATCTCAAGGCGGATTAACATGCGG 10099
QY 5101 CGTTTGGCGGTGTGATCCCTTTGATGAGGAGCGCAATGCTTCTGACAGTGA 5160
Db 10100 CGTTTGGCGGTGTGATCCCTTTGATGAGGAGCGCGCAATGCTTCTGACAGTGA 10159
QY 5161 GAACACACACTGATGAGGCGTACGTGAGTTCCTTCCAGATCATATGATCAGC 5220
Db 10160 GAACACACACTGATGAGGCGTACGTGAGTTCCTTCCAGATCATATGATCAGC 10219
QY 5221 AGTCGACCTAAAGTTCAACAGCTGCTCTGAAAGTCGCGCTGATATGATACGGAA 5280
Db 10220 AGTCGACCTAAAGTTCAACAGCTGCTCTGAAAGTCGCGCTGATATGATACGGAA 10279
QY 5281 CACCACTGCGGACCTGATAGCTTTGTCAATGCGTCAAGCGGATTCCTCAAGGACCT 5340
Db 10280 CACCACTGCGGACCTGATAGCTTTGTCAATGCGTCAAGCGGATTCCTCAAGGACCT 10339
QY 5341 GAAGTCATAGCAGGCGCGATATCAGCGGCTTTTACCCCTTTGACATTAAGTGTCTAT 5400
Db 10340 GAAGTCATAGCAGGCGCGATATCAGCGGCTTTTACCCCTTTGACATTAAGTGTCTAT 10399
QY 5401 CAGAAAGGGGCTTTGTTCACTACAGCTCCCTGATGAGAGTATGAAACAGAGAGC 5460
Db 10400 CAGAAAGGGGCTTTGTTCACTACAGCTCCCTGATGAGAGTATGAAACAGAGAGC 10459
QY 5461 GTTCGGCGATATTCAGGATCCTGCTTATGCTACAGACATAGTACCGGACCTGACAT 5520
Db 10460 GTTCGGCGATATTCAGGATCCTGCTTATGCTACAGACATAGTACCGGACCTGACAT 10519
QY 5521 ACGGCTGTGAAGCTTTCTGTCAAGAACATCAGCTCCCTTACACCCAGCAGATATCAGG 5580
Db 10520 ACGGCTGTGAAGCTTTCTGTCAAGAACATCAGCTCCCTTACACCCAGCAGATATCAGG 10579
QY 5581 GTATGAATGTGAAGAACACTCAGAGCAGCCCTGCAAGAAACAGACCAATTTGATG 5640
Db 10580 GTATGAATGTGAAGAACACTCAGAGCAGCCCTGCAAGAAACAGACCAATTTGATG 10639
QY 5641 TAAATTTGAAGTGAAGCTCTGCGAGCGCTTAAGTGTCTTACCGGACATCCCTATCTC 5700
Db 10640 TAAATTTGAAGTGAAGCTCTGCGAGCGCTTAAGTGTCTTACCGGACATCCCTATCTC 10699
QY 5701 GATTGACATCCCTGATGAGCTTTTGTAGATCATCAGAAATCACCACAAATTTTGAAGT 5760
Db 10700 GATTGACATCCCTGATGAGCTTTTGTAGATCATCAGAAATCACCACAAATTTTGAAGT 10759
QY 5761 TAGTCGACATAGAGATGATCATTTATTCGAGACATTTGAGTGTCTTCAATTAACA 5820
Db 10760 TAGTCGACATAGAGATGATCATTTATTCGAGACATTTGAGTGTCTTCAATTAACA 10819
QY 5821 GTACAAAGCTGACAGGAGGACATTTGTCAAGTTCATCCACTCCAGACAGAGCTGTTTT 5880
Db 10820 GTACAAAGCTGACAGGAGGACATTTGTCAAGTTCATCCACTCCAGACAGAGCTGTTTT 10879
QY 5881 GAAGGAAGCGACACATGTGATGCGGTGAGGACATTAACATTAATTTTACATC 5940
Db 10880 GAAGGAAGCGACACATGTGATGCGGTGAGGACATTAACATTAATTTTACATC 10939
QY 5941 GAGCCACACAGCAATTTATAGTTTGGCTATGCGGCAAGAGTCCACCTGCAATGCTGA 6000
Db 10940 GAGCCACACAGCAATTTATAGTTTGGCTATGCGGCAAGAGTCCACCTGCAATGCTGA 10999
QY 6001 ATGTAAACACCGCGGACACATATATTGAGAAACCATTAAGTGCACCAAGATTCCA 6060

Db 11000 ATGTAAACACCGCGGACACATATATTGAGAAACCATTAAGTGCACCAAGATTCCA 11059
QY 6061 GCGGCGACTTTCCAAACATCTTGGAACTGGCTGCTTGCATCTGTTGGGGAGCATCATC 6120
Db 11060 GCGGCGACTTTCCAAACATCTTGGAACTGGCTGCTTGCATCTGTTGGGGAGCATCATC 11119
QY 6121 CCTCATTTGTAGACTTATAGTGTGTGCTGACGCTCTATGCTTATTAACACAGCTAG 6180
Db 11120 CCTCATTTGTAGACTTATAGTGTGTGCTGACGCTCTATGCTTATTAACACAGCTAG 11179
QY 6181 ATGACTGACGCGGACACTGACATAGCGGTAAACCTGATGATCTTCCAGAAACGCTG 6240
Db 11180 ATGACTGACGCGGACACTGACATAGCGGTAAACCTGATGATCTTCCAGAAACGCTG 11239
QY 6241 TGCATATGCGACGCGCGCTTGGACATTAACCTGATGATTTTCCAGAAACAGCAGT 6300
Db 11240 TGCATATGCGACGCGCGCTTGGACATTAACCTGATGATTTTCCAGAAACAGCAGT 11299
QY 6301 CATATGCTGTCAGTGTACATTAATGCTATATCACTATCACTATTAACACACTATAT 6360
Db 11300 CATATGCTGTCAGTGTACATTAATGCTATATCACTATCACTATTAACACACTATAT 11359
QY 6361 CACTTTATGAGACTCACTATGAGTCTTATATATACATACATATTTTACTTAAAAAC 6420
Db 11360 CACTTTATGAGACTCACTATGAGTCTTATATATACATACATATTTTACTTAAAAAC 11419
QY 6421 ACTATACACACTTTATTAATTTCTTTTATATTTTCTTTTATTTTATTTTAA 6480
Db 11420 ACTATACACACTTTATTAATTTCTTTTATATTTTCTTTTATTTTATTTTAA 11479
QY 6481 ATTTC 6485
Db 11480 ATTTC 11484

RESULT 2
EEMSP
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
FEATURES
source
CDS
CDS

11470 bp ss-RNA linear VRL 02-AUG-1993
Western equine encephalitis virus RNA encoding structural proteins
at the 3' end of the genome.
J03854
J03854.1 GI:323728
6 kd protein; capsid protein; envelope protein.
Western equine encephalomyelitis virus
Western equine encephalomyelitis virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
Alphavirus; MEV complex.
1 (bases 1 to 4170)
Hahn,C.S., Lustig,S., Strouse,E.G. and Strouse,J.H.
Western equine encephalitis virus is a recombinant virus
Proc. Natl. Acad. Sci. U.S.A. 85 (16), 5997-6001 (1988)
88320369
3413072
Original source text: Western equine encephalitis virus (strain
BFS1703) RNA, passed in suckling mice and VERO cells, originally
from Cx. tarsalis.
Location/Qualifiers
1..4170
/organism="Western equine encephalomyelitis virus"
/mol_type="genomic RNA"
/db_xref="taxon:11039"
<1..114
/note="nonstructural protein 4"
/codon_start=1
/protein_id="AA42998.1"
/db_xref="GI:323729"
/translation="SRVEITLGLIISLSTLAEVKNFKSIRGNPITLVG"
159..3869
/note="envelope polyprotein"
/codon_start=1
/protein_id="AA42999.1"
/db_xref="GI:323730"

OY 3936 ACAGGCCAGGCCACACGCGTAAAGTCTATCTGAGAGAACGTCAGGCGAAGTGTCA 3995
DB 1621 ACAGGCCAGGCCACACGCGTAAAGTCTATCTGAGAGAACGTCAGGCGAAGTGTCA 1680
OY 3996 TTTAAACCACTTCTGCGAAGAACTGCTACCTACGAATGTAGTGGCGACTACAGCAG 4055
DB 1681 TTTAAACCACTTCTGCGAAGAACTGCTACCTACGAATGTAGTGGCGACTACAGCAG 1740
OY 4086 GTATGATGAGCAGCGGAAAGAAAGTAAAGCGGTGCACTTAAAGCAAAACAGTGCATTGCTT 4115
DB 1741 GTATGATGAGCAGCGGAAAGAAAGTAAAGCGGTGCACTTAAAGCAAAACAGTGCATTGCTT 1800
OY 4116 ACAAGAGGAGCCAAACGAATAGGATCTTCACTGCGCGGATCTTATTTAGGACACAGACC 4175
DB 1801 ACAAGAGGAGCCAAACGAATAGGATCTTCACTGCGCGGATCTTATTTAGGACACAGACC 1860
OY 4176 ACTGATGCAAGTAAATTTGCACTTCCATTCGCTTGAACACGACAGTCTGCGGTTT 4235
DB 1861 ACTGATGCAAGTAAATTTGCACTTCCATTCGCTTGAACACGACAGTCTGCGGTTT 1920
OY 4236 CGTTAGCTCACACGCGCTACAGTCAAGTGGTTTCAAGGCACTCACTTCCACTGACTG 4295
DB 1921 CGTTAGCTCACACGCGCTACAGTCAAGTGGTTTCAAGGCACTCACTTCCACTGACTG 1980
OY 4296 CAATGCGACCAACATTTGCTGACAAAGAAATTTGGGCTGAGACAGCAAGCAGCAG 4355
DB 1981 CAATGCGACCAACATTTGCTGACAAAGAAATTTGGGCTGAGACAGCAAGCAGCAG 2040
OY 4356 AATGATTAACAAGGCTCTACATCAGAGAAATTTTCTGTGGGGGAGAAAGGCTGAGTACG 4415
DB 2041 AATGATTAACAAGGCTCTACATCAGAGAAATTTTCTGTGGGGGAGAAAGGCTGAGTACG 2100
OY 4416 TATGGGATTAACATTAACCAAGTCAAGTCTGGGCCCAAGAGTCCGACCAAGGCCAC 4475
DB 2101 TATGGGATTAACATTAACCAAGTCAAGTCTGGGCCCAAGAGTCCGACCAAGGCCAC 2160
OY 4476 ATGATGCGCGCATGAGATCATCATCTCACTATTTATCATCGGCACTCAGTCTCACTGCA 4535
DB 2161 ATGATGCGCGCATGAGATCATCATCTCACTATTTATCATCGGCACTCAGTCTCACTGCA 2220
OY 4536 TTGTGCTGTGTGTGTGCTCTTGTCTATCTGTAGGCACTGCACTCATCAGACGTTGCA 4595
DB 2221 TTGTGCTGTGTGTGTGCTCTTGTCTATCTGTAGGCACTGCACTCATCAGACGTTGCA 2280
OY 4596 TCGCCAAAGCAAGAAAGAACTGCGCTGAGCGCACTAGCGCTTGCACCGAAGCGATAC 4655
DB 2281 TCGCCAAAGCAAGAAAGAACTGCGCTGAGCGCACTAGCGCTTGCACCGAAGCGATAC 2340
OY 4656 CCACAGCATTAAGCGGTTTGTGCTGATTCGCGCAACCAAGCGTGAACATTTGGAGAA 4715
DB 2341 CCACAGCATTAAGCGGTTTGTGCTGATTCGCGCAACCAAGCGTGAACATTTGGAGAA 2400
OY 4716 CTTTGAACCATCTGTGTTTAAACAACCGTTTCTGTGGCAGACAGTTGTGATCTCTC 4775
DB 2401 CTTTGAACCATCTGTGTTTAAACAACCGTTTCTGTGGCAGACAGTTGTGATCTCTC 2460
OY 4776 TGGCAGCGCTTGTATTTCTGTTCGCTGCTTTTCACTGCTGCACTGCTTTTATTTGTTG 4835
DB 2461 TGGCAGCGCTTGTATTTCTGTTCGCTGCTTTTCACTGCTGCACTGCTTTTATTTGTTG 2520
OY 4836 CAGGCGTCTGCGTGGGAGAGTGAAGCGCTTGAACATGCGACCACTGCGCAAAATGTTT 4895
DB 2521 CAGGCGTCTGCGTGGGAGAGTGAAGCGCTTGAACATGCGACCACTGCGCAAAATGTTT 2580
OY 4896 CCGGATCCCGTATTAAGCGTTGTGCAACGCGCAGTTTACCGGCACTTAACTTGAGGA 4955
DB 2581 CCGGATCCCGTATTAAGCGTTGTGCAACGCGCAGTTTACCGGCACTTAACTTGAGGA 2640
OY 4956 TCAAGGTGTCTCATCGGAATTAACAACCTTCACTTAAACAAGAGTACGTGACCTGCAAT 5015
DB 2641 TCAAGGTGTCTCATCGGAATTAACAACCTTCACTTAAACAAGAGTACGTGACCTGCAAT 2700

OY 5016 TCCACACAGTCACTTCTTCAACCAAGTTAAATGCTGCGGTCCTCGATGCAAGCAT 5075
DB 2701 TCCACACAGTCACTTCTTCAACCAAGTTAAATGCTGCGGTCCTCGATGTAAGCAT 2760
OY 5076 CCTCAAAGCGGATTTACACATGCGCGCTTTTGGCGGTGTGACCTTTTATGTGGGAG 5135
DB 2761 CCTCAAAGCGGATTTACACATGCGCGCTTTTGGCGGTGTGACCTTTTATGTGGGAG 2820
OY 5136 GCGCAATGCTTCTGTGACAGTGAACACACAATGAGGAGCGTACCGTCAAGTTG 5195
DB 2821 GCGCAATGCTTCTGTGACAGTGAACACACAATGAGGAGCATACCGTCAAGTTG 2880
OY 5196 CTCACAGTGAATATGATACCGACGTGCACTTAAAGTTTCAACACAGCTGCTGGAAG 5255
DB 2881 CTCACAGTGAATATGATACCGACGTGCACTTAAAGTTTCAACACAGCTGCTGGAAG 2940
OY 5256 TCGGCTGCGTATGATATGCGCAACACACCGCGCACTTGGATACGTTTGTCAATGCG 5315
DB 2941 TCGGCTGCGTATGATATGCGCAACACACCGCGCACTTGGATACGTTTGTCAATGCG 3000
OY 5316 TCAAGCGAGTTCCTCAAGGAGCCTGAAGGTCATAGAGGCGCGATACAGCGGCTTTT 5375
DB 3001 TCAAGCGAGTTCCTCAAGGAGCCTGAAGGTCATAGAGGCGCGATACAGCGGCTTTT 3060
OY 5376 CACCTTTGACCATTAAGGTGCTCATCAGAAAGGCGCTTTTCACTACAGACTTCCCTG 5435
DB 3061 CACCTTTGACCATTAAGGTGCTCATCAGAAAGGCGCTTTTCACTACAGACTTCCCTG 3120
OY 5436 AGTATGAGCTATGAAACAGAGAGGTTTGGCGATATTCAGACATCTTCTGATGCTA 5495
DB 3121 AGTATGAGCTATGAAACAGAGAGGTTTGGCGATATTCAGACATCTTCTGATGCTA 3180
OY 5496 CAGACATGTAAGCCCGCACTGACATACGCGCTGCTGAGACCTTCTGTAAGAACATCCAC 5555
DB 3181 CAGACATGTAAGCCCGCACTGACATACGCGCTGCTGAGACCTTCTGTAAGAACATCCAC 3240
OY 5556 TCCCTTCAACCAAGCAGTATCAGGCTATGAAATGTGAAGAACCACTCAGAGCAGCCC 5615
DB 3241 TCCCTTCAACCAAGCAGTATCAGGCTATGAAATGTGAAGAACCACTCAGAGCAGCCC 3300
OY 5616 TGCAGAAACAGCAACATTTGATGTAATTTGAAGTGAAGCTTCTGAGCGTTTCACT 5675
DB 3301 TGCAGAAACAGCAACATTTGATGTAATTTGAAGTGAAGCTTCTGAGCGTTTCACT 3360
OY 5676 GTGCTTACGGGCAATCCCTATCTGATGACATCCCTGAGACGCTTTTGTGATAT 5735
DB 3361 GTGCTTACGGGCAATCCCTATCTGATGACATCCCTGAGCGCTTTTGTGATAT 3420
OY 5736 CAGAAATCAACAATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 5795
DB 3421 CAGAAATCAACAATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3480
OY 5796 ACTTTGATGTTCTTAAACATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 5855
DB 3481 ACTTTGATGTTCTTAAACATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3540
OY 5856 ACTGCCACTCCAGCAGAGCTGTTTGAAGAAAGGAGCAACACATGTGATGCTGAGGCA 5915
DB 3541 ACTGCCACTCCAGCAGAGCTGTTTGAAGAAAGGAGCAACACATGTGATGCTGAGGCA 3600
OY 5916 GCATTAACATTAATTTTGAAGCATGAGCCCAAGCAAGCAATTTTATGTTTGGCTATGCG 5975
DB 3601 GCATTAACATTAATTTTGAAGCATGAGCCCAAGCAAGCAATTTTATGTTTGGCTATGCG 3660
OY 5976 GCAAGAACTCCAGCTGCAATGCTGAATGTAAACAACCGCGCAGCACTAAATTTGAGAAC 6035
DB 3661 GCAAGAACTCCAGCTGCAATGCTGAATGTAAACAACCGCGCAGCACTAAATTTGAGAAC 3720
OY 6036 CACATAAAGTGAACCAAGAAATTTCAAGGCGGCAAGTTTCCAAACATCTTGAATGCTGCG 6095
DB 3721 CACATAAAGTGAACCAAGAAATTTCAAGGCGGCAAGTTTCCAAACATCTTGAATGCTGCG 3780
OY 6096 TTGCACTGTTTGGGAGACATATCCCTCATTTGTTGATGAGCACTATATGTTGTTGCTGCA 6155

```
Db 3781 TTGACGTTGGGGGAGCATCATCCCGCATTGTTGTAAGACTTATAGTGTGGTCGCA 3840
Qy 6156 GCTTATAGCTTATTAACAACACAGTATGATGACTGAGCGGAGACACTGACATAGCGGTAAC 6215
Db 3841 GCTTATAGCTTATTAACAACACAGTATGATGACTGAGCGGAGACACTGACATAGCGGTAAC 3900
Qy 6216 TCGATGACTTCCGAGGAGCGTGTGATTAATGCAACGCGCGCTTGACACTAAACTC 6275
Db 3901 TCGATGACTTCCGAGGAGCGTGTGATTAATGCAACGCGCGCTTGACACTAAACTC 3960
Qy 6276 GATGATATTTCCGAGGAGCAACAGTCAATATGCTGTGAGTGCATTAATCGTATATC 6335
Db 3961 GATGATATTTCCGAGGAGCAACAGTCAATATGCTGTGAGTGCATTAATCGTATATC 4020
Qy 6336 ACACCTACATATTAACAACACTATATCACTTTATGAGACTCAGATGGGCTCTTAATATA 6395
Db 4021 ACACCTACATATTAACAACACTATATCACTTTATGAGACTCAGATGGGCTCTTAATATA 4080
Qy 6396 CACTACACATATTTACTTAAACACATATACACTTAAATCTTAAATCTTATAATTTT 6455
Db 4081 CACTACACATATTTACTTAAACACATATACACTTAAATCTTAAATCTTATAATTTT 4140
Qy 6456 CTTTGTGTTTATTTGTTTAAATTTT 6485
Db 4141 CTTTGTGTTTATTTGTTTAAATTTT 4170

RESULT 3
AF229608 4047 bp RNA linear VRL 27-MAR-2000
LOCUS Western equine encephalomyelitis virus structural polyprotein gene,
DEFINITION complete cds.
ACCESSION AF229608
VERSION AF229608.1 GI:7330240
KEYWORDS Western equine encephalomyelitis virus
SOURCE Western equine encephalomyelitis virus
ORGANISM Alphavirus, WEEV complex.
REFERENCE 1 (bases 40 to 816)
AUTHORS Uryvaev,L.V., Lebedev,A.Iu., Sokolova,T.M. and Inferov,V.P.
TITLE Primary structure of the nucleocapsid gene C and the protein coded
by it from the Western equine encephalomyelitis virus
JOURNAL Dokl. Akad. Nauk. 344 (3), 397-401 (1995)
MEDLINE 96097271
PUBMED 8535272
REFERENCE 2 (bases 40 to 816)
AUTHORS Uryvaev,L.V. and Lebedev,A.Iu.
TITLE Comparative analysis of primary structure of nucleocapsid protein
from Western equine encephalomyelitis virus and other alphaviruses
JOURNAL Vopr. Virusol. 41 (6), 252-259 (1996)
MEDLINE 97130212
PUBMED 8999681
REFERENCE 3 (bases 997 to 2265)
AUTHORS Uryvaev,L.V., Lebedev,A.Iu., Sokolova,T.M. and Inferov,V.P.
TITLE Capsid glycoprotein of B2 Encephalitis virus, Western Equine:
PRIMARY structure of gene and its product
JOURNAL Dokl. Akad. Nauk. 357 (1), 134-139 (1997)
MEDLINE 9471275
PUBMED
REFERENCE 4 (bases 1 to 4047)
AUTHORS Uryvaev,L.V.
TITLE Primary structure of WEEV 26S RNA
JOURNAL Unpublished
REFERENCE 5 (bases 1 to 4047)
AUTHORS Uryvaev,L.V.
TITLE Direct Submission
JOURNAL Submitted (29-JUN-2000) Molecular Biology, D. I. Ivanovsky
Institute of Virology, 16 Gamaleya str., Moscow, Russia
FEATURES
Source 1..4047
/organism="Western equine encephalomyelitis virus"
```

```
CDs
/mol_type="genomic RNA"
/strain="McMillan"
/db_xref="taxon:11039"
40..3747
/codon_start=1
/product="structural polyprotein"
/protein_id="AF60166.1"
/db_xref="GI:7330241"
/translation="MPPYPOLNPPVYPTNPMAYRDPNPPRCMRPRPPLAIOIEDL
RRIITLITFKQRPANPPPCPPPKKSSAPPRKPTQKKKKQAKTKRKPKPKRQRM
CMRLSDKTPPLPLNGVNGVACVAGRLMKPLHLHGKLDNEQLAKLKKASVYDLS
YGVAVPNNMSDITLQYTSDKPPGYNHMGVQYENRFTVPRGVGKSGSRIPLDK
GRVAVLPGANGRTALSVYMNQKGLVITKDPGSEBWSLVTAICVSNVTFPD
KPPVCSLAPERTDLVLENDVNPNTDILENVLCPSRPKSIITDDFLTSPYLG
CPNCRSAPCPSPRIKIENWDESDDSIRVSAOYGYNAGTADTKPPMYMDH
DIEDSMEKAIATSGPCRLGKGYFLAQCRRGSVYTSIGSASENCTYEFKR
RKFGREBYLFPVPHKRYVCHVYDHLKERSAYITMHRSPHAYSVYLEASGEVIT
KPPSGKNVYIECKGDSYTGIVSTPKMNGCTAKOCIAKSDQTKWVNSPDLIRHT
DHSVQKHLHPRLIPVCPPLAPLPTVTKMFGJITLHLATRPILLTRKGLRAD
ATSEWITGTSRNSFVGRGLEEYVMGNHEPVRVAOESAPSDHGMPEHIIHYHRH
PVYTVVLGVALAILVGTASSACTTKARDCLTPYALA PMATVPTALALCIRPT
MATPFGTLLNHLFNNQPLMAQLCTPLAALLIFRCFSCEMFLVAGCLGVDAF
EHAFTVNPVGIPIKALVERAGYAPLNLEIKVSSBELTSTNKEVYTCFHYVPSFO
VKCCGLECKASSKADYTCRFGVAYPEFMWGACQCEDSENTQLSEAYFEAPCTTD
HAVALRHPTALRVGLRIYVGNNTAYLDTFVNGVDTGSSRLVIAGPISAARSFPH
KVIVIRGLVYNDPPEFGAMKPGAFGDIQASSDADIVARTDIRLLKPSVKNIPHY
TOAVSGYEMKNNSGRPLEBTAPGCKIEVPEPRASNCAYGHIPISIDIPDAFVRS
ESPFLIEVSTADCIYSADPGSLTQYKADRGCHCPYHSHSTAYULKEATHTYAG
SITLHFSTSSPQANFIVSLDCGKTTCAECKPPADHIIIGEPHYVDPEFOAAVSKTSM
MILALFPGASSLIVGLIYVCSMLINTR"
40..816
/product="C-protein"
/note="nucleocapsid protein"
/citation={1}
/citation={2}
817..996
/product="E3-polypeptide"
/peptide="E2-protein"
/note="envelope glycoprotein"
/citation={3}
2266..2460
/product="6K-polypeptide"
/peptide="E1-protein"
/note="envelope glycoprotein"
2461..3744
/product="E1-protein"
/note="envelope glycoprotein"
3748..4047
BASS COUNT 1125 a 1041 c 962 g 919 t
ORIGIN
Query Match 59.4%; Score 3850.2; DB 14; Length 4047;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 3939; Conservative 0; Mismatches 108; Indels 4; Gaps 2;
Qy 2435 ATAGTGACGTAGTAGACAGCACTTACCCAGCGGAGAAATGTTTCATACCTCACTG 2494
Db 1 ATAGTGACGTAGTAGACAGCACTTACCCAGCGGAGAAATGTTTCATACCTCACTG 60
Qy 2495 AACTTCCACCAAGTTTACCTTACCAATCGATGGCTTACCGAGATCCAAACCTCTAGG 2554
Db 61 AACTTCCACCAAGTTTACCTTACCAATCGATGGCTTACCGAGATCCAAACCTCTAGG 120
Qy 2555 CGCCGCTGAGAGCGCTTGGCCCGCGCTGCTCAATTCGAAGTCTTAGAGGTCG 2614
Db 121 TGGCGCTGAGAGCGCTTGGCCCGCGCTGCTCAATTCGAAGTCTTAGAGGTCG 180
Qy 2615 ATAGTCACTTGACTTCAACAACAGATCACTTAATCCGCGCAGGCTCCACGCGCAAG 2674
Db 181 ATAGTCACTTGACTTCAACAACAGATCACTTAATCCGCGCAGGCTCCACGCGCAAG 240
Qy 2675 AAGAAGAAAGTGCTCTTAAGCAAAACCTACTGAGCTTAAAAAGAAAGAAAGCAAGCC 2734
Db 241 AAGAAGAAAGTGCTCTTAAGCAAAACCTACTGAGCTTAAAAAGAAAGAAAGCAAGCC 300
```


QY 2735 AAGAGACGAAACGACGCTTAAACAGAGAAACGACAGTATGTATGAAAGTTGGAG 2794
Db 301 AAGAGACGAAACGACGCTTAAACAGTGAAGAAACGACAGCTATGTATGAAAGTTGGAG 360
QY 2795 TCGGACAGACATTTCCGATCATGCTGAACGGCCAAAGTAATGATATGCTGCTTGTCT 2854
Db 361 TCGGACAGACATTTCCGATCATGCTGAACGGCCAAAGTAATGATATGCTGCTTGTCT 420
QY 2855 GGAGGAGGCTGATGAACCACTCCAGCTTGAAGGAAAAATTGATTAATGAGCAATTAACG 2914
Db 421 GGAGGAGGCTGATGAACCACTCCAGCTTGAAGGAAAAATTGATTAATGAGCAATTAACG 480
QY 2915 GCCGTAATTAAGAAAGGCTAGCATGATGACGCTTGAAGTACGAGGCTTCCAGAAC 2974
Db 481 GCCGTAATTAAGAAAGGCTAGCATGATGACGCTTGAAGTATGAGGAGCTTCCAGAAC 540
QY 2975 ATGAATTAAGACACGCTGATGACACGAGCGCAACCAACCGGCTTCTACAACTGGAC 3034
Db 541 ATGAATTAAGACACGCTTCAATGACACGAGCGCAACCAACCGGCTTCTACAACTGGAC 600
QY 3035 CACGCGAGCTCCAGTATGAGATGGGAGATTACCGTACCGAGAGAGTGGCGGAGAA 3094
Db 601 CACGCGAGCTCCAGTATGAGATGGGAGATTACCGTACCGAGAGTGGCGGAGAA 660
QY 3095 GGCGACAGCGGAAAGACGATCCTGGAACAGAGGAGAGTGTGGCTATTGTTCTAGAA 3154
Db 661 GGCGACAGTGAAGACGATCCTTACACAGAGGACAGTGTGGCTATTGTTCTAGAA 720
QY 3155 GGTGCAATTAAGAGGACCGCTTCAAGGCTTTCAGTGTCTCACTTGAACCAAGAGGGGTG 3214
Db 721 GGTGCAATTAAGAGGACCGCTTCAAGGCTTTCAGTGTCTCACTTGAACCAAGAGGGGTG 780
QY 3215 ACCATTAGGATACCCCGGAAGTCTTGAACCGTGTGTCACTAGTATGACGCTATGCTG 3274
Db 781 ACCATTAGGATACCCCGGAAGTCTTGAACCGTGTGTCACTAGTATGACGCTATGCTG 840
QY 3275 CTTTGAATGTCACTTCCCATGCGACAAACCAACCGTGTGTCTCACTAGTATGCTG 3334
Db 841 CTTTGAATGTCACTTCCCATGCGACAAACCAACCGTGTGTCTCACTAGTATGCTG 900
QY 3335 CGAACACTGACGCTGTGGAAGAGAGCTGCAACATTCMAATTACGACACGCTGTGAG 3394
Db 901 CGAACACTGACGCTGTGGAAGAGAGCTGCAACATTCMAATTACGACACGCTGTGAG 960
QY 3395 AACGCTTGAATGTCCATCAACGCGCGCCAAACCAACATTCGAGATTCACACTG 3454
Db 961 AACGCTTGAATGTCCATCAACGCGCGCCAAACCAACATTCGAGATTCACACTG 1020
QY 3455 ACCAGTCCCTACCTGGGCTTCTGACCGTATTGACAGACTCAACGCGCTTTCAGCCA 3514
Db 1021 ACCAGTCCCTACCTGGGCTTCTGACCGGAAATGACAGACTCAACGCGCTTTCAGCCA 1080
QY 3515 ATAAAAATTGAAGAGCTGTGGAAGAGATCTGATGATGATGATGATGATGATGATGATG 3574
Db 1081 ATAAAAATTGAAGAGCTGTGGAAGAGATCTGATGATGATGATGATGATGATGATGATG 1140
QY 3575 GCACAATTCGGCTACATCAAGGACGCTGCGGATGTCACCAATTCCTGTTACATGCT 3634
Db 1141 GCACAATTCGGCTACATCAAGGACGCTGCGGATGTCACCAATTCCTGTTACATGCT 1200
QY 3635 TTGACACGACGATGATCAAGGAGAGAGATGAGGAAAAATTGATTAATGAGCAATTC 3694
Db 1201 TTGACACGACGATGATCAAGGAGAGAGATGAGGAAAAATTGATTAATGAGCAATTC 1260
QY 3695 GGAAGCTGCGCTGTCTTGGCCAAAGGATATCTCTGTTAGCTCAATGTCCTCAAGT 3754
Db 1261 GGAAGCTGCGCTGTCTTGGCCAAAGGATATCTCTGTTAGCTCAATGTCCTCAAGT 1320
QY 3755 GACAGTGAACCGCTAGATACAGAGCGGAGAGATTTGAATTCATGACCGTGGAGAA 3814
Db 1321 GACAGTGAACCGCTAGATACAGAGCGGAGAGATTTGAATTCATGACCGTGGAGTT 1380

QY 3815 AAGATCAGAGAGAGTTTGTGCTAGAGAGAGTACTTGTCCACCCGCTCATGAAAG 3874
Db 1381 AAGATCAGAGAGAGTTTGTGCTAGAGAGAGTACTTGTCCACCCGCTCATGAAAG 1440
QY 3875 CTGGTAAAGTCCACGTTTACATCACTTGAAGAGAGCTGCGGGTACATTAACATG 3934
Db 1441 AGGGTAAAGTCCACGTTTACATCACTTGAAGAGAGAGTCTGCGGGTACATTAACATG 1500
QY 3935 CACAGGCGGACCAACAGGCTATGATGCTATGCTGAGAGAGAGTCTGAGGAGAGTATG 3994
Db 1501 CACAGGCGGACCAACAGGCTATGATGCTATGCTGAGAGAGAGTCTGAGGAGAGTATG 1560
QY 3995 ATTAACCACTTCTGACAGAGAGTCACTTACAGATGATGATGATGATGATGATGATG 4054
Db 1561 ATTAACCACTTCTGACAGAGAGTCACTTACAGATGATGATGATGATGATGATGATG 1620
QY 4055 GGTATCGTGAACGCGGACAGAGATGAGAGGCTGCACTTAAAGCAAAACAGTCAATGCC 4114
Db 1621 GGTATCGTGAACGCGGACAGAGATGAGAGGCTGCACTTAAAGCAAAACAGTCAATGCC 1680
QY 4115 TACAAGAGCGACCAAGGAAATGGTCTTCACTCCGCGATCTTATTAAGCAGACAGAC 4174
Db 1681 TACAAGAGCGACCAAGGAAATGGTCTTCACTCCGCGATCTTATTAAGCAGACAGAC 1740
QY 4175 CACTCAGTGAAGGTAATTAATGACATTCATTCGCTTGAACAGCAGTCTGCGCGTT 4234
Db 1741 CACTCAGTGAAGGTAATTAATGACATTCATTCGCTTGAATTAAGCAGTCTGCGCGTT 1800
QY 4235 CCGTTAGCTCAACGCTTACAGTCAAGAGTGTCAAGGAGTCAACCTCACTGACT 4294
Db 1801 CCGTTAGCTCAACGCTTACAGTCAAGAGTGTCAAGGAGTCAACCTCACTGACT 1860
QY 4295 GCAATGCGACCAATGCTGACAAAGAGAAATTTGGGCTGCGAGCAGAGCAGCAACAGCA 4354
Db 1861 GCAATGCGACCAATGCTGACAAAGAGAAATTTGGGCTGCGAGCAGAGCAGCAACAGCA 1920
QY 4355 GAATGATTAAGAGGCTTCACTCAGAGAAATTTCTGCGGGGCGAGAGGCTGAGATAC 4414
Db 1921 GAATGATTAAGAGGCTTCACTCAGAGAAATTTCTGCGGGGCGAGAGGCTGAGATAC 1980
QY 4415 GTATGGGTTAACAATGAACAGTCAAGTCTGGGCGGAGAGTGGGCAACAGGCGACCA 4474
Db 1981 GTATGGGTTAACAATGAACAGTCAAGTCTGGGCGGAGAGTGGGCAACAGGCGACCA 2040
QY 4475 CATGATGCGCGGATGATGATCATCATCACTTATATATGCGGATTCAGTCTTACATGCT 4534
Db 2041 CATGATGCGCGGATGATGATCATCATCACTTATATATGCGGATTCAGTCTTACATGCT 2100
QY 4535 ATTGCTGTGTGTGCTGCTCTTGTATCTGTGAGGACATGATCATCAGAGCTTGC 4594
Db 2101 ATTGCTGTGTGTGCTGCTCTTGTATCTGTGAGGACATGATCATCAGAGCTTGC 2160
QY 4595 ATGCGCAAGAGAGAGAGCTGCGCTGACGCGCATACGCGCTTGAACCGAACGCAACGCTA 4654
Db 2161 ATGCGCAAGAGAGAGAGCTGCGCTGACGCGCATACGCGCTTGAACCGAACGCAACGCTA 2220
QY 4655 CCCACAGATTAAGAGGCTTGTGCTGATTCGCGCAACCAACGCTGAACATTTGAGAA 4714
Db 2221 CCCACAGATTAAGAGGCTTGTGCTGATTCGCGCAACCAACGCTGAACATTTGAGAA 2280
QY 4715 ACTTTGAACCATCTGTGTTTAAACAACCAACGTTTCTGAGGACAGTGTGATGCTCT 4774
Db 2281 ACTTTGAACCATCTGTGTTTAAACAACCAACGTTTCTGAGGACAGTGTGATGCTCT 2340
QY 4775 CTGACAGGCTTGTATTTCTGCTGCTGCTTTCATGCTGATGCTTTCATTTATGCTT 4834
Db 2341 CTGACAGGCTTGTATTTCTGCTGCTGCTTTCATGCTGATGCTTTCATTTATGCTT 2400
QY 4835 GCAAGCGCTGCTGCGGAGAGTGAAGCGCTTGAACATGAGACATCTGTCGCAAAATTT 4894
Db 2401 GCAAGCGCTGCTGCGGAGAGTGAAGCGCTTGAACATGAGACATCTGTCGCAAAATTT 2460
QY 4895 CCGGAGATCCGCTATTAAGAGCGTGTGTCGAACGCGCAGGTTACGCGCACTTAACCTGAG 4954

```

Db      2461 CCGGGATCCCGTATAGGGCGTTGTCGAGTCGAGGTTACGGCCACTTAATCTGAG 2520
Qy      4955 ATCAGGTCGTCTATGGAATTAACACCTTCAACAAGGATACCTGCAAA 5014
Db      2521 ATTAAGTGTCTATGGAATTAACACCTTCAACAAGGATACCTGCAAA 2580
Qy      5015 TTCACACAGTCATTCCTTCCACCAAGTTAAATGCTGGGGTCCCTCGAGTGCAGGCA 5074
Db      2581 TTCACACAGTCATTCCTTCCACCAAGTTAAATGCTGGGGTCCCTCGAGTGCAGGCA 2640
Qy      5075 TCCTCAAGGCGGATTAACAATGCGCGGTTTGGCGGTGTGAACCTTTGATGCGGGA 5134
Db      2641 TCCTCAAGGCGGATTAACAATGCGCGGTTTGGCGGTGTGAACCTTTGATGCGGGA 2700
Qy      5135 GGGGCAACATGCTTCTGTGACAGTGAACAACAATGAGTGGGGTGTGAGTTC 5194
Db      2701 GGGGCAACATGCTTCTGTGACAGTGAACAACAATGAGTGGGGTGTGAGTTC 2760
Qy      5195 GCTCCAGACTGCACTATAGATCAAGCACTGCGCACTAAAGTTCACAGCTGCTGAAA 5254
Db      2761 GCTCCAGACTGCACTATAGATCAAGTGCAGTGCAGTTCACACAGCTGCTGAAA 2820
Qy      5255 GTGGGCTGGTATAGTATACGCGAACAACAACCGCGACCTGATAGCTTTGTCAATGGC 5314
Db      2821 GTGGGCTGGTATAGTATACGCGAACAACAACCGCGATCTGATAGCTTCATATGGC 2880
Qy      5315 GTACCGGCAAGTTCCTACCGGGACCTGAAGTCTATAGCAGGGCGGATACGCGCTTTT 5374
Db      2881 GTACCGGCAAGTTCCTACCGGGACCTGAAGTCTATAGCAGGGCGGATACGCGCTTTT 2940
Qy      5375 TCACCTTTTGACATTAAGTGTGTCATCAGAAAGGGGCTGTTTCAACTACGACTCCCT 5434
Db      2941 TCACCTTTTGACATTAAGTGTGTCATCAGAAAGGGGCTGTTTCAACTACGACTCCCT 3000
Qy      5435 GAGTATGAGATGATAAACAAGAGCGTTCGCGATATTCAGACATCTCGCTTGATGCT 5494
Db      3001 GAGTATGAGATGATAAACAAGAGCGTTCGCGATATTCAGACATCTCGCTTGATGCT 3060
Qy      5495 ACAGACATGATAGCCCGGACCTGACATACGCTGCTGAAGCTTCTGTCAACAATCAGC 5554
Db      3061 ACAGACATGATAGCCCGGACCTGACATACGCTGCTGAAGCTTCTGTCAACAATCAGC 3120
Qy      5555 GTCCCTTACACCCAGGATGATGAGGATGATGAAGTGAAGCAACTCGACGACGCC 5614
Db      3121 GTCCCTTACACCCAGGATGATGAGGATGATGAAGTGAAGCAACTCGACGACGCC 3180
Qy      5615 CTGCAAGAAACAGCACCATTGATGATAAAATTGAAGTGAAGCTCTGCGACGCTTAAC 5674
Db      3181 CTGCAAGAAACAGCACCATTGATGATAAAATTGAAGTGAAGCTCTGCGACGCTTAAC 3240
Qy      5675 TGTCCTTAAGGGGACATCCCTATCTGATTAACATCCCTGATGAGCTTTGTAGATCA 5734
Db      3241 TGTCCTTAAGGGGACATCCCTATCTGATTAACATCCCTGATGAGCTTTGTAGATCA 3300
Qy      5735 TCAGATACACCAACAATTTTGAAGTTAGTGCACAGTAGACAGCTGACTTTATTCGCA 5794
Db      3301 TCAGATACACCAACAATTTTGAAGTTAGTGCACAGTAGACAGCTGACTTTATTCGCA 3357
Qy      5795 GACTTTGGTGTCTCTAACATTAAGTCAAGTCAAGGAGGAGCAATGTCAGATT 5854
Db      3358 GACTTTGGTGTCTCTAACATTAAGTCAAGTCAAGGAGGAGCAATGTCAGATT 3417
Qy      5855 CACTCCCACTTCACGACAGCTGTTTGAAGAGGAGCACACATGTCAGCTGAGGC 5914
Db      3418 CACTCCCACTTCACGACAGCTGTTTGAAGAGGAGCACACATGTCAGCTGAGGC 3477
Qy      5915 AGCATTAACACTAATTTTATAGACATGAGCCACAAAGCAATTTTATGTTGCTATGC 5974
Db      3478 AGCATTAACACTAATTTTATAGACATGAGCCACAAAGCAATTTTATGTTGCTATGC 3537
Qy      5975 GCGAAGAGTCCACCTGCAATGCTGAATGTAACCAACCGCGGACCACTAATTTGAGAA 6034

```

```

Db      3538 GCGAAGAACCAACCTGCAATGCTGAATGAACAACCGGCGCACACATTAATTTGAGAA 3597
Qy      6035 CCACATAAGTCGACCAAGAAATTCAGGGCGGAGTTTCCAAACATCTTGAAGTGGCTG 6094
Db      3598 CCACATAAGTCGACCAAGAAATTCAGGGCGGAGTTTCCAAACATCTTGAAGTGGCTG 3657
Qy      6095 CTTCGACTGTTTGGGAGGAGCATCCCTCATTTGTTGTAGACTTTATAGTGTGCTGC 6154
Db      3658 CTTCGACTGTTTGGGAGGAGCATCCCTCATTTGTTGTAGACTTTATAGTGTGCTGC 3717
Qy      6155 AGCTCTATGCTTTAATAACACAGTATGACTAGAGCGCGGACACTGACATAGCGGTAAA 6214
Db      3718 AGCTCTATGCTTTAATAACACAGTATGACTAGAGCGCGGACACTGACATAGCGGTAAA 3777
Qy      6215 CTGATGATCTTCCGAGGAAGCGGTGTCATTAATGACGCGCGCTTGACACTAAACT 6274
Db      3778 CTGATGATCTTCCGAGGAAGCGGTGTCATTAATGACGCGCGCTTGACACTAAACT 3837
Qy      6275 CGATGATTTTCCGAGGAAGCAGATGATGCTGTGACAGTGCATTAATGATAT 6334
Db      3838 CGATGATTTTCCGAGGAAGCAGATGATGCTGTGACAGTGCATTAATGATAT 3897
Qy      6335 CACACTACATATTAAACAACATATATCATTTTATGAGACTCACTATGGGCTTAATAT 6394
Db      3898 CACACTACATATTAAACAACATATATCATTTTATGAGACTCA-TATGGGCTCTAATAT 3956
Qy      6395 ACCTACACATATTAACTTAATAAACAATATACACTTTATTAATCTTTTATTAATTT 6454
Db      3957 ACCTACACATATTAACTTAATAAACAATATACACTTTTATTAATCTTTTATTAATTT 4016
Qy      6455 TCCTTTGTTTATTTTATTTTAAATTTTC 6485
Db      4017 TCCTTTGTTTATTTTATTTTAAATTTTC 4047

RESULT 4
AF339476
LOCUS
DEFINITION
Highlands J virus strain B-230 nonstructural protein 4 gene,
partial cds; and structural polyprotein precursor, gene, complete
cds.
ACCESSION
AF339476
AF339476.1 GI:28193935
VERSION
AF339476
KEYWORDS
ORGANISM
Highlands J virus
Highlands J virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
Alphavirus; WEV complex.
REFERENCE
1 (bases 1 to 4692)
Kinney,R.M. and Pfeiffer,M.
Nucleotide sequence analyses of the 26S mRNAs of viruses of the
genus Alphavirus
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 4692)
Kinney,R.M. and Pfeiffer,M.
Direct Submision
TITLE
Submitted (23-JUN-2001) Arbovirus Diseases Branch, Centers for
Disease Control and Prevention, P.O. Box 2087, Fort Collins, CO
80522, USA
FEATURES
source
1..4692
location/Qualifiers
/organism="Highlands J virus"
/virus
/mol_type="genomic RNA"
/strain="B-230"
/db_xref="taxon:11024"
/country="USA"
/note="isolated in 1960; virus passage P5/SW3/BHK21-1"
<1..564
/codon_start=1
/product="nonstructural protein 4"
/protein_id="A03332.1"
/db_xref="GI:28193936"

```

/translation="FKFGAMMKSGNFLLFVNTLVNIMIASRVLERLTSSACAFIG
DNIIVHGVSDNIMADRCATWLMNEKLIIDAVIGKAYCFGLVVOIGIACRVA
DPLKRLFKLGLPLDDODVDRRALDESLRMRIISITSELVAVESREVLVLSL
ITIALSTLADSVKNKIRGNPITLYG"
565..4692
/product="structural polypeptide precursor"
/note="268"
565..605
606..4316
/codon_start=1
/product="structural polypeptide precursor"
/protein_id="AA033323.1"
/db_xref="GI:28193937"
/translation="MFYPOLTFPMYQPNMAYRDPNPPRRMRPRLAOTIEDL
RRIANLTKORANPPAPPAKKKKTKPKAPPKKKKKOAVKKKKRPPKRROR
CMKLESDTFPILNGQVNGYACVGVGLMRHLYEGLIDNEQLAILKXAMWLE
YGDVPONKSDTLQYTESKPEFTNMHGAQYENGRTVPRGVGKSDSRPLDNK
GRVAIVLKGANESRTALSVTNQKQVITKIDPEGEQMSLTAMCVLANVPPD
KRPVCSLAPRSLDLEENYNDPCYDILNVLKCPERROKRSITDFTLTSPLGH
CPYCLHATPSPSIKIENWDESDGTIRIQVSAQGNQAGTADPTFRVMSYEDH
DIKESMDKIAISTGSCSLGHKGIFILARCPDSDVTSGASATSCVEKRI
RPFVGREBYLLPYHNGKIKCHVDHLETTAGYITMRPGRHATVYESSEKVI
RPPGRNVTYRECKSDISTGTNTRTKMPEGCKKQCIAYRHQDKVFNPSDLIRHS
DHAVQKLIIFKLTAIACVPLATPTVEKMFKQVTLALASHPTLLTRKLGPRAR
PTSEWIVGTVRNFSVREGLYETWGNMPPRVWSOBAPDPMHPEIIVHYHRH
PLTYIIVLGLVLTIVIGIASAAICISARBDCLTPYALAPNAVPTLAVLCIRPT
HAETLGESLGHMLNNOPLMAOLCLPLAAILIFRFSCLPILVAGTADGADAY
EHATTVPVPGVPKALYERAGYAPLNEVTVSESLPSTNKEVYTCRHTIIPSP
VKCCGSLCQASRKADITCRVFGVYPPRMWGAQCFCSENTQJSEAVEPAPDCTAD
HAVALKVTAAIKVGLQIVGNTSTRIDTFVNGVTPGAGTISAEVISAAPFPD
KVIRKGVNIVPEPEGMKRGVGLDQASLSDSTIDVARDTLIDLSVKSIIHVPY
TOAAGYEMKNNNSGRPLQDTAPFGCKLEVPDLRAVDCAVGHIPLSIDIIPAAFVRS
EAPVLEMSCTVCTACTVADFGSLTLQYKDKBNCVHSHSAUVLEKATTVVHS
GSVTLHSTSSPQANFIYSLCGKKTCAECKPESDILIGEPHKVNGEFOAAVSKTSH
NWLPMLGASSLIVGLVLVACSSMIINTSR"
mat_peptide
/product="capsid protein"
606..1382
1383..1442
/product="E3 protein"
mat_peptide
/product="E2 envelope glycoprotein"
1443..2831
2832..2996
/product="E6 protein"
mat_peptide
2997..4313
/product="E1 envelope glycoprotein"
4317..4692
3'UTR
BASE COUNT 1324 a 1183 c 1105 g 1080 t
ORIGIN

Query Match 41.7% Score 2705.6; DB 14; Length 4692;
Best Local Similarity 75.6%; Pred. No. 0; Mismatches 1084; Indels 4; Gaps 2;
Matches 3380; Conservative 0;

QY 1866 TTTAAATTGGTGCATGATGAAATCCGATATGTTCTTAACGCTGTTTGTCAACACACTA 1925
DB 1 TTTAAATTGGTGCATGATGAAATCCGATATGTTTGAAGCTGTTTGTGAATATCTCTA 60
QY 1926 GTCAATATCATGATTGCTAGCAGAGTACTAGCTGAACGTTTAAACAGCTCAGCGCCG 1985
DB 61 GTTAAACATCATGATGACAGTGCAGTACTAGAGAAAGTTGACCATCTCAGCGTGTCA 120
QY 1986 GCTCTATCGGAGCATTAACATAGTGCATGCTGCTCCGACACCTTAATGCGGAG 2045
DB 121 GCTCTTATCGAGACGACATATATGATACCGGTGATGATCCGATTAATTTGATGCTAT 180
QY 2046 AGATGCGGCACTGGCTGAACATGAGAGTAAATATATGATGATGATGATGATGATGAT 2105
DB 181 AGGTGTGCGACGTGCTAAATATGAGAGTAAATATGATGATGATGATGATGATGATGAT 240
QY 2106 GCAACCTTCTTCTGTGGGGGATTTATCTGTGTGACCGAGTAAACGACGACGCTGCA 2165
DB 241 GCACCATCTTTTGGGAGGGGTTCACTCTGTGTGATCAAAATTAACAGGTGATGAG 300

QY 2166 GTCCGAGACCCCTTAATAAAGGCTTTTAAAGCTGGAAAAACATTCGCAGTCATGATACC 2225
DB 301 GTAGAGAGACCCCTTAACACGTTTGTTTAACTGGGTAAACACTCCCTTTAGACATGAT 360
QY 2226 CAAGCTGCGACCCGCGCGGCACTGCATATGAAAGCAATGCGATGGAACGAATTGCA 2285
DB 361 CAGATGTAGACCCCGCGGCGGCTGACATGATGCTTAAGGTGGAATATGATAGT 420
QY 2286 ATTACGAGCGATTAAGTAAAGCCGTGAAATCCAGATATCATCTGCGAGCCCTG 2345
DB 421 ATACCTCTGTGACTTTGTTAAAGCGGTGAAATCCGATATGAGGTGCTCTGTAGCTTA 480
QY 2346 ATCATCAGCTCTGTGCTCACCTTAAGCCGAAGACCTTAAGAACTTAAGAGCATAGAGG 2405
DB 481 ATTATTTATGGCTGTGCAACTTGCATATGCGTGAAGAACTTTAAACCGATTAAGAGA 540
QY 2406 AGCCCAATCACCTCTTACGCGTGAACCTTAATAGTGAACGTATGACACGACCTAACCA 2465
DB 541 AACCCATTACTCTTTAGGCTGACCTTAATAGGTGACGTATGATACGCACTA---A 597
QY 2466 CCGGAGAAATGTTTCCATACCTCAGCTGAACCTTTCCACAGTTTACCTTAACCAATCCGA 2525
DB 598 CCGGAAATATGTTCTTATCCGCAATTGACTTTTCCACTATGTATGAACCCAACTCTA 657
QY 2526 TGGCTTACCGAGATCCAAACCTCTAGGCGCGCGTGAAGCCGTTTGGCGCCCGCTGG 2585
DB 658 TGGCATACAGAGACCCGAACCTCTTAGCGCCAGTGAAGCCATTAGGCCCCCACTTG 717
QY 2586 CTGCTCAATGGAAGATTAGAGAGTGCATATGTCACCTTGACCTTCAACACGATCAC 2645
DB 718 CAGCACAAAATAGAGACTTGAAGAGTGCATCGCAAAATTTGACCTTAAACAGAGAGAC 777
QY 2646 CTAATCCGCCCGCAGTCCACCGCCAAAGAAAGAAAGAGTCTCTTAAGCCAAACCTA 2705
DB 778 CCAATCCACTGCGCGTCTCTCACTTAAGAAAGAAAGAAACCTCAACCGAAACCGAAGGCA 837
QY 2706 CTCAGCCTTAATAAAGAAAGAGAGCAAGGCAAGAGCAAGCAAGCTTAACCAAGGGA 2765
DB 838 CGCCACCTTAAGAAAG 897
QY 2766 AACGACAACTGATGTATGTAAGATTGAGAGTGGAGCAAGACATTTCCGATCATGCTGAA 2825
DB 898 AAGGCAACGATGTATGTAAGATTGAGAGTGGAGCAAGACATTTCCGATCATGCTGAA 957
QY 2826 GCCAAGTAAATGATATGCTGCTGCTGCTGCGAGAGAGGCTGATGAACCACTCCACGTTG 2885
DB 958 GGCAGGTAATGAGCTACGCAATGCTGCTGCGAGAGGCTATGAAACATTTGATGATGCTG 1017
QY 2886 AAGGAAAAATTGATATGAGCAATTAAGCGCGCGTGAATTGAAGAGGCTGACATGTAG 2945
DB 1018 AAGGAAAAATTGATATGAGCAATTAAGCGCGCATCACTGAAGAAAGGCGATGTAGCG 1077
QY 2946 ACTTGAGTACGAGGAGCGTTTCCGAGAACATGAATCAGACAGCTGACATACACAGCG 3005
DB 1078 ATTCGAGTACGAGGAGCGTTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1137
QY 3006 ACAAACCAACCGGCTTCTGACAACTGSCACACCGGCGCAGTCCAGTATGAAATGGAAT 3065
DB 1138 AAAAAACCACTGATATCTATTAATCTGCAATCTGCGCAGTCCAGTATGAAATGGAAT 1197
QY 3066 TTACCGTACGAGAGAGAGTGGGCGGGAAGGCGACAGCGGAAGACCGATCTCTGACACA 3125
DB 1198 TCACCGTTCACGCGGAGTGGTGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1257
QY 3126 GAGGACAGATTGTGCTATTTCTTGAAGAGTGCAGATGAGGCAACGCTGACGCGCTTT 3185
DB 1258 AAGGCGGAGTGTGCTATTTCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1317
QY 3186 CAGTGTCTACTTGGAAACGAAAGGAGTGCATTTAGGAGATACCCCGGAAGCTTTGAA 3245
DB 1318 CGTGTGTTCATGTGAACCAAAAGGAGTGCATTCAGAGATACCCCTGGAAGGCAACGAA 1377
QY 3246 CGTGTCTACTGTGAACGCGCTATGCGTGTGCTTGAATGTACAGTTCCATGCGCAAAAC 3305

|||||
Db 1378 AATGTCATTGATCACTGSCAATGTCGTCCTTGCAAAATGTCATCTTCCCGTGACAAAC 1437
Qy 3306 CACCGGTGCTGATTTCACTGACGCCAAGAACCACTGACGTGCTCGAAGAAAGCTCG 3365
Db 1438 CTCCTGTCTGCTATTCGTAGCACCAGAACGCACACTTGATGTCTGGAAAGAAACCTGG 1497
Qy 3366 ACAAATCAAAATTAGACACGCGTGTGAGAAAGCTTGAAATGTCATACGCGCGGCCA 3425
Db 1498 ACAAATCAAGTTAGACACCTGCTAGAAATGTAATGAAATGCGCCCTCACGTGACAGA 1557
Qy 3426 AACGAACATTACCGATGACTTCACTGACAGTCCCTACCTGCGGTTCTGCCGTAAT 3485
Db 1558 AAGGACATTTAGGAGAGATTTCACTTGACAGACCGGTAATTAGGCGCATGTCGTAAT 1617
Qy 3486 GCAGACACTCAACCCGCTGTTTCCAGCCCAATAAAAATTGAGACGTGTGGAGCAATCTG 3545
Db 1618 GCCCTCACGCTACCCCTGCTTTAGTCCATTTAAATTGAAACGTTTGGGACGAGTCAG 1677
Qy 3546 ATGATGATGATTAAGAAATCCAGGTCGCGCAATTCGGCTACATTCAGGACAGGACCTG 3605
Db 1678 ACGACGTAACAATGAAATACAGTTTCGCCAATTAAGTTACAAACGACGAGAGTACGG 1737
Qy 3606 CGGATGTCACCAAAATTCGGTTACATGTCCTTTCGACACAGACCATGACATCAAGAGACA 3665
Db 1738 CAGACCAACTAATGATCAGGTACATGTCCTACAGACCAACATGACATCAAGAAAGCTA 1797
Qy 3666 GTATGAGAAAATATGCTATCAACACATCTGGAACCTCGCGCTGCTTGGCCACAAAGGT 3725
Db 1798 GTATGAGCAAGATGCTATCAATGTCATCTGGGCGGTGACGCGCTTGGGCGATTAAGCT 1857
Qy 3726 ACTTCCTGTTAGCTCAATGTCCTCAGGTGACAGGTGAACCGTACAGTATCAAGAGCGAG 3785
Db 1858 ATTTCTGCTGGCCAGATGCCACACAGAGACAGCGTAAACGTACATCAACAGTGGAG 1917
Qy 3786 CATCTGAAATTCATGACCCGTGAGAAAAGATCAGAGAGAGATTTGTCGTTAGAGAG 3845
Db 1918 CATAGGAGACATCTGACGCGTGCAGAAAGAAATAGAGAGAAATGTCGCGGAGAGAG 1977
Qy 3846 AGTACTGTTCCACCCGTCATGAAAGCTGTAAAGTGCACAGTTTACATCACTTGA 3905
Db 1978 AGTACTGCTCCACAGTGACGCGCAAGCTAATCAAGTGTCAATGATATACCACTAA 2037
Qy 3906 AGGAGAGTCTGCGGGGTACATAACATGACAGGCGACGCGCAACGCGATTAAGTCT 3965
Db 2038 AGGAGACAAAGAGATATATATATGACACAGGCGCAACATGCGTAACCAACT 2097
Qy 3966 ATCTGAGAGAGCGTCAAGGCGAAGTGTACATTAACCACTTCTGCGAAGACGTCACT 4025
Db 2098 ACGTGAAGAAATCATCGGCGAAGGTATCATTCGTCACCTTCGGAAAGAACGTGACGT 2157
Qy 4026 ACGAATGTAAGTGGCGACTACAGCACAGGATTCGTGAGACCGCAAGCAAGATGAACG 4085
Db 2158 ACGAATCAATGTGAGATTACAGCACCGGACACGTTAACATAGAACTTAAGATGCGAG 2217
Qy 4086 GCTGACATTAAGCAAAACAGTGCATTTGCTCAAGAGCGACCAACGAATGGGTCTCA 4145
Db 2218 GATCAGAGAAAGAAACAGTGCATAGCTTTAAACAGACCAACCAATGGGTCTTTA 2277
Qy 4146 ACTGCGCGATCTTATTAAGGACACAGACCACTGACAGTGAATTAATTGACATTCAT 4205
Db 2278 ATTACACAGACTTAATTCGTACCTCAGACCATGACAGTGCAGAAACTGACATTCCT 2337
Qy 4206 TCCGCTTGACACGACAGTCTGCGCGGTTCCGTTAGCTCAGACGCTACAGTCAAGT 4265
Db 2338 TTAAACTTAACGTCTACGCGCATGCCAGTGCACCTCGCTACACACCAACGCTTAAGAAAT 2397
Qy 4266 GGTTCAAAGGATACCTCCACTGACTGCAATGCGACCAACATGCTGCAACGAGAA 4325
Db 2398 GGTTCAAAGGCTACGCTGACCTGACTGCTAGTCAATCCACACTGTTGACTCGGCTA 2457
Qy 4326 AATGGGCTGCGAGACGCAACAGCAGATGAGATTACAGGCTCTCATCCAGAAAT 4385
|||||

Db 2458 AGCTAGGACCGGAGCTGAACCCACGTCGATGATCTGTTGGAGCCGTCGAGAAACT 2517
Qy 4386 TTTCTGTGGGGCGAAGAGGCTGAGTACGTATGGGTTACATGAACCGTCAAGTCT 4445
Db 2518 TCTCCGAGGAGAGAGAGGCTCGAATATATCTGGGAAATTCAGACCCGTCGAGTGT 2577
Qy 4446 GGGCCGAGAGTGGGACACAGGCGACCCACATGGAATGAGCGCATGATCATCCACT 4505
Db 2578 GGTCTCAAGAGTCTGACCTGGGAGCCCTCACGCTTGCGGACAGAAATATGTCATAT 2637
Qy 4506 ATTATCATGGGATTCACGTCATACCTGTCAATGCTGTGTGTGTGCTCTTGTCTATTC 4565
Db 2638 ACTACATCGGCAATCCGCTATACCAATTCGCTGTGTGTCGCACTGTCTAATAACAG 2697
Qy 4566 TGGTAGGCACTGATCATACGAGCTTGATCCGCAAGAGCAAGAGAGATGCTGACAC 4625
Db 2698 TGATGGCATGCTTTCGACGCGGCTGCAATATGAAAGCGGAGAGATGCTGACAC 2757
Qy 4626 CATACGCTTGACCGAAGCAAGGATCCACAGCATTAAGGCGTTTGTGTCATTC 4685
Db 2758 CTAAGCGTTGGACCCACAGCTGACATTCGAATGTCAGTCTTATGTCGATTA 2817
Qy 4686 GGCACCAACCGCTGAACATTTGAGAACTTTGAACATCTGTGTTTAAACCAAC 4745
Db 2818 GACCTACGACGCGGAGACGCTAGGCGAATCACTTGACACCTCGCTTAACATCAAC 2877
Qy 4746 CGTTCTCTGGGACAGTTGTCATCTCTGACAGGCGTTGTTATCTGTCCGCT 4805
Db 2878 CTCCTTTATGGGCGACGTGCTTCACCTTGCGCATTAATCATCTTTCAGTTC 2937
Qy 4806 TTTCATGCTCATGCTTTTATTTTATGTTGTCAGGCGCTGCTGCGGAAAGTACGCT 4865
Db 2938 TCTCATGCTGCTCTCTTTTATTTGTTGACGCGCTGCTGCGGAAAGGACGCT 2997
Qy 4866 TCGAATGCGACCATCTGTCGCAAAATGTTCCGGAGATCCGTAATAGCGCTTGTCAG 4925
Db 2998 ACGAATGCGACCATCTGTCGCAAAATGTTCCGGAGATCCGTAATAGCGCTTGTCAG 3057
Qy 4926 GCGGAGTTACGGGCACTTAACCTGAGATCAACGTCGCTCATGCGAATTAACCTT 4985
Db 3058 GGGCGGGATGACCCCTCAACCTGAGATCAACGTTGTGTGTCGATGATACCT 3117
Qy 4986 CAATTAACAAGAGTACGTACCTGCAAAATTCACACAGTCAATCTTCCACCAAGTTA 5045
Db 3118 CAACAAATTAAGAAATATGATGATCTTGAATTCATTAATCAATCCACAGGTGA 3177
Qy 5046 AATGCTCGGCTCTCTGAGTGAAGGATCTCAAAAGCGGATTAACATGCCGCTT 5105
Db 3178 AGTGTGCGGTTCCCTTGAATGTCAGGCTCACGCAAGGACGACTATACCTGTGCGCTGT 3237
Qy 5106 TTTGCGGTGTATCCCTTATGATGGGAGGCGCAATGCTTCTGAGACAGTGAACA 5165
Db 3238 TCGAGGTGTATCCCAATTAATGAGGAGGCGCCAGTCTTCTGAGATGCGAANAACA 3297
Qy 5166 CACAATGATGAGGCTGATGAGTTCGCTCAGACTGACATATAGATCAAGCACTG 5225
Db 3298 CGCAATGAGAGAGGATGATGATGATTCGCTCAGACTGACAGCAAGATCAAGCACTG 3357
Qy 5226 CACTAAAGTTTACACAGCTGCTGAAAGTGGGCTGCTGATATATATGCGCAACCA 5285
Db 3358 CACTGAAGTTCACACAGCTGCTTGAAGGTCGGCTGCAAAATATATATGAGAAACAT 3417
Qy 5286 CCGGCACTGATACCTTTGTAAATGCGCTCACGCGCAAGTTCCTCACGCGAAGCTG 5345
Db 3418 CAACACGCTGAGATCAATTTGTCAATGAGGATTAACACAGGCTCATACGAGACTG 3477
Qy 5346 TCAATGAGGAGGATTAATGACGCTTTTACCTTTGACCAATAGAGTGTCTACGAA 5405
Db 3478 TGATGAGAGGCTTATCTACAGAGCTTACACCTTTGACACCAAGTGTCTATTCGAA 3537
Qy 5406 AGGGGCTGTTTAACTACGACTTCCCTGATATGAGAGTATGAAACAGAGAGCTTGC 5465
Db 3538 AAGGTAAAGTATCAATTAAGACTTCCAGAGTACGAGTGTATGAAAGCAGGCGTGTTCG 3597
|||||

QY 5466 GCGATATTCAGCATCTCTGCTTATGCTACAGACATAGTAGCCCGACATGACATACGGC 5525
Db 3598 GGGATATTCAGGCAATCTTCTGCTGATTCACAGACATGTTGCAAGAACTACATCCGAC 3657
QY 5526 TGTGGAAGCCCTTCTGTCAAGAACATCCAGTCCTCCCTACACCCAGCAAGTATCGGGTATG 5585
Db 3558 TACTCAAAACCATCGGTAAAGAGCATTCACGTACCGTACACAGAGCCCGTCCAGTTACG 3717
QY 5586 AATGTGGAAGAACACTCAGAGACCCCTGCAAGAAAGACACCATTTGGATGTAAA 5645
Db 3718 AATGTGGAAGAAATATTCGGAAGACCTCTGACAGATACAGCCCTTTTGCTGCAAAA 3777
QY 5646 TTGAAGTGAGACCTCTGCGAGCCGTCTAATCTGTCTTACGGGACATCCCTATCTCATATG 5705
Db 3778 TCGAGTAGATCCGTTAAGGGCGGTTGACTGTGCTTACGGCCATATCCCGCTCCATCG 3837
QY 5706 ACATCCCTGATGACGCTTTTGTAGATCATGAATCCAAACAAATTTTAAAGTTAGCT 5765
Db 3838 ACATACCGGAGCGAGCATTCGTGAGAACCTCTGAAGACCCACCGTCTCGAAATGAGTT 3897
QY 5766 GCACAGTAGAGACATCTTATCTGAGACTTTGGGTCTTCTAATTAAGTATCA 5825
Db 3898 GCACGGTGAACGCTCTGACATTTACCTCCGATTTTGGATCTTTAACTCCACAAATCA 3957
QY 5826 AAGCTGACAGGAGGAGCATTTGTCAGTTCACTCCCACTCCAGACAGCTGTTTGAAG 5885
Db 3958 AAGGAGACAAAGAAAGGCAATTTGCTGTCTGTCATTTGCACTCTCTACGGCGGTGCTGAAG 4017
QY 5886 AAGGACCAACATGATGATCGCTGAGGACGATPAACATTAATTTAGACATCGAGCC 5945
Db 4018 AAGTACACACCCAGTGTACTCGGGAGCGTTACTTTACCTTGACGACGTCAGTC 4077
QY 5946 CACAAGAAATTTATAGTTTCTGATGCGGCAAGAAAGTCCACTGCAATCTGAATGTA 6005
Db 4078 CGCAGGCAAACTTATTTGCTGCTATGCGGTAAAGAAACAAATGCGACGCGAGTGA 4137
QY 6006 AACCAACGCGGACGACATTAATTTGAGAACACATTAAGTGCACCAAGAAATTTCCAGCGG 6065
Db 4138 AGCCCCATGACACATCATAGGGGAAACACATTAAGTAAACCAAGAAATTTCCAGCAG 4197
QY 6066 CAGTTTCCAAAACATCTTGAACCTGGCTGCTTGCACGCTTTGGGGAGCATCATCTCA 6125
Db 4198 CAGTTTCCAAAACATCATGAACTGGCTATTCGCGCATGCTGGGGAGCATCATCTCA 4257
QY 6126 TTGTTGAGAGCTATATGTTGTTGCTGACCTTATGCTTATTAACACAGTATGATGAC 6185
Db 4258 TTGTTGAGAGCTATATGTTGTTGCTGACCTTATGATTAATCAACACAGTATGATGAC 4317
QY 6186 TGAACGCGGACACTGACATAGCGGTA-AACTGATGACTTCCGAGGAAGCGATGCA 6244
Db 4318 CCGGCGTTCCTCTACACAGCGGCAAGCACTCAATGACTTCCGAGGTAACGATGCA 4377
QY 6245 TAATGCAACGCGCGCTTGAACATAATGATGATTTTCCGAGGAAGCAAGTGCATA 6304
Db 4378 TAATGCAACGCGCGCTTGAACATAATGATGATTTTCCGAGGAAGCAAGTGCATA 4437
QY 6305 ATGCTGTGCACTGTCATTAATGAT 6332
Db 4438 ATGCTGTGCACTGTCATTAATGAT 4465

RESULT 5
AR042411
LOCUS AR042411 11663 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 7 from patent US 5811407.
ACCESSION AR042411
VERSION AR042411.1 GI:5962907
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 11663)

AUTHORS Johnston, R.E., Davis, N.L., and Simpson, D.A.
TITLE System for the in vivo delivery and expression of heterologous
JOURNAL genes in the bone marrow
Patent: US 5811407-A 7 22-SEP-1998;
FEATURES
Source Location/Qualifiers
1..11663
BASE COUNT 3289 a 3105 c 2910 g 2359 t
ORIGIN
Query Match 37.0%; Score 2397.6; DB 6; Length 11663;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3752; Conservative 0; Mismatches 1984; Indels 74; Gaps 9;
ORIGIN
QY 580 ACTGACGTTATGAAGCGGAGCGGTATATTTCTCATGGAAGAGCCCAAGTACCTTC 639
Db 5707 ACTGTTAACCAGGGGTAGGTGGTACATATTTTGCAGGACACAGGCCCTGGGCACTTGC 5766
QY 640 AACGAATGACATGCTCAATGTAATACACAAAGACCTATATTTGATCGGCCCTGCATG 699
Db 5767 AAAAGAGTCCGTTCTGACAGACAGCTTACAGAACCGACTTGGAGCGCAATGTTCTG 5826
QY 700 AGAAGTTACGCCCCCGCTCGATCTCGAAGAGAGAAATGTACAGAAAGACTGC 759
Db 5827 AAGAAATCTACGCCCCGCTCGATCTCGAAGAGAGAAACAGCTCAATCAGTACC 5886
QY 760 AATATCGCTCTGAAGGAATTAAGACAGTATCAATCAAGAAAGTAAATATGA 819
Db 5887 AGATGATGCCACCGAAGCCAAAGACAGTACAGTCTGAAAGTAAAGAACCA 5946
QY 820 AAGCAATTAAGCGGAGCGACTCATTTCTGATTTGGGACACATCTATCATCAAGATGA 879
Db 5947 AAGCATTAACCATGACGACGACTGTTTCAAGGCTACAGCTGTATTAATCTGCCACAG--- 6003
QY 880 ATCTGTGAGTTTACAGATCAATTAATCTGTAACATCTGTAACAGGTATTA 939
Db 6004 ATCAAGCAAGATGCTATTAAGATCACTACCCGAACCATGCTATTCAGACAGTATACAG 6063
QY 940 ACAGTTTACATGCGAGGTCGGGTAAACGTCACATTAATGTTATCCAAAGATTT 999
Db 6064 CGAATCTCTGACCCAAAGTTTGTAGCTGTTTGTATCAACTATCTGATAGAAATTT 6123
QY 1000 ACCCTACAGTAGCCAGTTATTTGTAACAGATGAATACGATGCTATTTGACATGCTG 1059
Db 6124 ACCCGAGGTAGCATCTTATCAATACACGACAGTACGATGCTTATGATATGATG 6183
QY 1060 ACGGCGATCGTGTCTGTAGATACAGCCACTTTTTCGGCTAAACAGAAAGTACC 1119
Db 6184 ACGGAGCAGTCGCTTGTGCTGATGATCTCAACTTTTTCGCCGCAAGCTTGAAGTTACC 6243
QY 1120 CAAAGAAAGCATGATTTTGGACGCGAGATAGATAGCCGTCCTTATGATGTA 1179
Db 6244 CGAAAGACACGATTAAGAGCCCAAAACATCCGCACTGGGTTCATACGCCGATGCA 6303
QY 1180 ATACATTACAAATATGTTTGGCTGACGCTACTAAAGAAATGCAAGTTACCCAAATGC 1239
Db 6304 ACAGTTTGAAGAGTCTCTATGTCGCGACTAAAGAAATGCAAGTCTACAAATATGC 6363
QY 1240 GAGATTAACCTGTTTGAATTTGCGCGCATTTATGTTGATTTGTTCAAGAAATAGCAT 1299
Db 6364 GTGAATCGCAACACGAGTACAGCGACATTTCAACGTTGATGCTTGAATAATATGAT 6423
QY 1300 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1359
Db 6424 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6483
QY 1360 TTACGCAATATGTCACAAAGCTGAAAGGCGGAAAGCAGCAGCTGTTTGGCAATATCTC 1419
Db 6484 TTACGCAATATGTCACAAAGCTGAAAGGCGGAAAGCAGCAGCTGTTTGGCAATATCTC 6543
QY 1420 ATATATCTAAACCGTTGACAGAGATACCAATGATCAATTTGTCATGATCTAAAGAG 1479
Db 6544 ATATATCTAAACCGTTGACAGAGATACCAATGATCAATTTGTCATGATCTAAAGAG 6603

QY 1480 ATGTCAGTAACTCCCGGACGAAACATACAGAGAGCGGCTTAAGTGGAGTTATTC 1539
DB 6604 ACCTGAAAGTTACCTCGGACGAAACACAGAGAAAGACGAAAGTTCAAGTGTAC 6663
QY 1540 AGGCTGAGATCCCTTGTCTACCGCTTACCTTGGGGATCCATCGGAAATTAGTCCGTA 1599
DB 6664 AAGCCGAGAAACCCCTGGCGCGCTTACCTATGCGGATCCACCGGAGTTAGTGGCA 6723
QY 1600 GACTGAATGGGCTCTTGGCAATATCCATCTCTTGGACATGTCAAGCGGAAGATT 1659
DB 6724 GGGCTTACGCGCTTGTCTACCCAACTTCAACGCTTGTGACATGTCCGCGAGACT 6783
QY 1660 TTGATGATTTATTTGTAACATTTCCACACCGGCGACCCAGTTGGAACCGACATCG 1719
DB 6784 TTGATGATTTATTTGTAACATTTCCACAGAGGTGACCGGTAATGGAAGCGATATCG 6843
QY 1720 CGTCTGTTGATTAAGGAGAGACGACGCTATCCGCTATTTGGCGCTTGATATCTTGAGG 1779
DB 6844 CTTCTGTTGATTAAGGAGAGACGACGCTATTTGGCTTACCGGCTGATGATCTTGAGG 6903
QY 1780 ACTTGGTGTGACCAACCGCTCTTGAATTTGATAGGCGCGCTTGGCAATATCAT 1839
DB 6904 ACCTGGTGTGACCAACCACTACTCGACTTGTGAGTGGCGCTTGGAAATATCAT 6963
QY 1840 CTGTGACCTACTACGAGAGAGGTTTAAATTTGGTGCATGATGAATCCGCTATGT 1899
DB 6964 CCACCCACTTGGCCACGAGTACCGCTTCAATTTGGGGGAGATATTAATCCGAAATGT 7023
QY 1900 TCTTAAAGCTGTTTGTCAACACATAGTCAATATCATGATTTGTGAGAGTACTAGTG 1959
DB 7024 TCTTCAAGCTTCTTGTCAACACAGTCTGAATGTCGTTATCGCAGAGATTTGAGG 7083
QY 1960 AAGGTTTAAACAGTCAAGCGTGGCGGCTCTATCGGCGACGATTAATAGTGTATGTG 2019
DB 7084 AGGGCTTAAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 7143
QY 2020 TCGTCTCCGACACCTTGAAGGAGATGCGGCACTGGCTGAAATGGAAGTAAATA 2079
DB 7144 TAGTATTTGCAAGAAATGCTGAGAGTGTGCTGCTCAATGGAAGTAAATA 7203
QY 2080 TTAATGATGATTAATTTGATTAAGACACCTTATCTGTGGGGATTTATCTGTGTG 2139
DB 7204 TCAATGACGAGTATGCGGAGAGACCACTTACTTCTGGGTGATTAATCTTGTGAG 7263
QY 2140 ACCAGATTAACAGGACAGCTGCAAGTCCGACACCTCTTAAAGGCTTTTAACTTG 2199
DB 7264 ATTCTGTTACTCTCAACGCTGTGCGTGGCGGACCTTGAAGGCTGTTAACTTGG 7233
QY 2200 GAAACCATTTGCAAGTGAATGATCCGAGCTGCGACCGCGCGGGCACTGATGATG 2259
DB 7324 GTTAAACGCTTCCGCGGAGGATGACAGACGAGACGAGCGGCTCTGTGATAGT 7383
QY 2260 AAGCAATGCAATGGAACAGATTTGAATTAAGGAGTGAAGGCGGTGAATCCA 2319
DB 7384 AAGCAATGGAAGGCTGTTAAGATGATTAAGACACCTTGAAGGCGGTGCAATCTC 7443
QY 2320 GATTAACGATTAATCTGCGAGGCTGATCAATCACTCTGTGCACTTGAAGCGGAAGCG 2379
DB 7444 GGTATGAGTGAACATCAACATCACTCTGTGCTGAGATTTGAAGCTTTGCGGAGCA 7503
QY 2380 TTAAAGACTTCAAGAGCAATGAAGGAGCCCAATCACTCTGAGGTGACTTAAATAGG 2439
DB 7504 AAGAGCATTTCAAGCATCAAGAGGGAATTAAGCATCTCTGAGGTGCTTAAATAGT 7563
QY 2440 TGAAGTATGA-----CAAGCACTTACCAACCGGAGAGATTTTCCA 2482
DB 7564 CAGATATGATTAATTTCACTGATTAATTAACCAACACCAACATGAATAGAGATTTCT 7623
QY 2483 TACCTTCAAGCTTCAACCACTTCAACCAATTCGATGCTTAAACGAGATCCA 2542
DB 7624 TTAACATGCTGCGCGCGGCTTCCAGCGGCGGCACTGCTGATGAGGCGCGGAGAA 7683

QY 2543 AACCTCTAGGCGCGCTGAGGCGGCTTTGCGGCGGCTGCTCTCAATCGAAGAT 2602
DB 7684 GAGGAGAGGCGCGCGGATGCTGCGGCAATGGCTGCTTCAATCGAAGATCGA 7743
QY 2603 CTGAGAGTGTGATGATCACTTCAATTAACAA--GATCACTTAATCCGCGCGAG 2660
DB 7744 CCACACCGCTGATGCTTGAATTTGAGACAGGCACTGACCTCAACCCCGGCGCAC 7803
QY 2661 GTCCACCGCAAGAAAGAAAGATGCTCTTAAGCAAACTACTACGCTTAAAGAA 2720
DB 7804 GCGGCGCGCGCGCGCAAGAAAGAGCGCAAGGAAACCAACCGGAGCGGAGAAACAA 7863
QY 2721 AGAAGAGCAAGGCAAGAGAGCAAGAAC---GCAAGCTTAAACGAGGAAACGAACAGTA 2777
DB 7864 AAGCAAGAGAAAGAAAGAAAGCAACCTGCAAAACCAACCCGGAAGAGACAGGTA 7923
QY 2778 TGTGATGAAGTTGAGTGTGAGCAAGACATTTCCGATC--ATGTAACGCGCAAGTGA 2834
DB 7924 TGGCACTTAAGTTGAGGCGGAGCAGACTGTTGACCTCAAAATGAGAGCGGAGATGTA 7983
QY 2835 ATGATATGCTGCTGCTGCTGCGAGGAGGCTGATGAACCACTCCAGTTGAGGAA 2894
DB 7984 TCGGAGACGACTGCGCATGAGAAAGATTAAGAAACCACTCAAGTAAAGAACTA 8043
QY 2895 TTGATATGAGCAATTAGCGGCGGTGAATTAAGAAAGGCTGATGATGATGAGT 2954
DB 8044 TTGACCACTGCTGCTGCTGATTAAGCTCAATTTCAACCAAGTGTGAGCATGATGAGT 8103
QY 2955 ACGGCGACGTTCCCGCAACATGAATCAGACAGCTGCACTACCAAGGCAACAAACAC 3014
DB 8104 TCGCAGAGTTGCGGCTGCAACATGAGAAAGTGAAGCGTTCACTACACAGTGAACACCTG 8163
QY 3015 CGGCGTTTCAACATGAGCAACGCGGCGAGTCCAGTATGAGAAATGGAAGTTAACGCTAC 3074
DB 8164 AAGGTTCTTCAACATGAGCAACGAGCGGTGAGTATGAGAGGACATTTTACATCC 8223
QY 3075 CGAGAGAGTGGCGGGAAGAGGAGACAGGAGAAACCGATCTTGAACACAGAGGAGAG 3134
DB 8224 CCGCGAGATGAGAGGAGAGGAGAGAGAGAGTGTGTCGATTAAGTAACTACAGGCGCGG 8283
QY 3135 TTGTGCTATTTGTTTGAAGAGTGCATTAAGGAGACGCTTACGCGCTTCACTGTGTA 3194
DB 8284 TTGTCCGATGATCTCGGAGGCTGATGAGAGGAAACAAAGACCGCTTTCGCTGTCA 8343
QY 3195 CTGGAACCAAGAAAGGAGTGCATTAAGGATACCCCGAAGGTTGGAACGCTGTG--- 3251
DB 8344 CTTGAATGAGAAAGGAGAAACATCAAGCAACCCCGAGGAGACAGAAAGATGCTGTG 8403
QY 3252 -----CACTAGTACGCGCTATGCGTCTTTCGATGTCAGTTCCTCATGCGAACAAC 3305
DB 8404 CTGACACCTGCTGACAGGCTATGCTTGTGGAAGCGTACCTTCCATGCAATCGGC 8463
QY 3306 CACCGGTGTCTATTTCACTGACGCGCAAGCAACTCGACGTGCTGGAAGAGAGTCTG 3365
DB 8464 CGCCCATGCTTACACCCCGGAAACATCAAGCTCTGCACTCTCGAAGAAAGTGA 8523
QY 3366 ACAATCAATTAAGCAACGCTGCTGGAAGAGTCTGGAAT---GTCAATCAAGCGCGG 3422
DB 8524 ACCAGAGGCTTACGACACCTGCTCAAGCGCAATTTGGGTGGATGCTCCGAGAA 8583
QY 3423 CGAAGAGATTAACGATGATGATCACTGACAGTCCCTTACTTGGGTTCTGCGCGT 3482
DB 8584 GTAAAGAGAGGCTGACGAGACTTTTACCTTGAACAGCCCGTACTTTGGGACATGCTCGT 8643
QY 3483 ATTGACAGACCTCAAGCGGCTTTTCAAGCCCAATTAATTTGAACGCTGTGGAGCAAT 3542
DB 8644 ACTGTACATATGTAACCGTGTCTTGAAGCGGCTTGAAGTGAAGTGTGAGTGAAG 8703
QY 3543 CTGATATGAGATGATTAAGATCAGGCTCGGCACAATTCGCTCAATCAAGGAGCA 3602
DB 8704 CGGAGCAACCAACATACGATACAGATTCGCGCCAGTTTGAATACGACAAAGCGGAG 8763
QY 3603 CTGCGATGTCACCAATTCGTTACATGCTTTGACACGACGACATGATCAAGGAAG 3662

Dp	8784	CAGCAGCTCAAAATAGTACCGCTACATGTGCTGACGAGATATATCTGTCAAGAG	8823
Qy	3663	ACAGTATGAGAAAATAGCTATCAGCACATCGACCCTGCGTCTTGACCAAAAG	3722
Dp	8824	GCACCATGGATGACATCAAGATCAGCACTTCAGGACCGGTATGAAAGCTTATGCTACAAAG	8883
Qy	3723	GGTACTTCTCTGTAGCTCAATGTCTCTCAAGGTGACAGTGTAACTCTACGTATCAGACCG	3782
Dp	8884	GATACTTCTCTCTCCGGAAGTGTCTCCAGGGGACAGGTAACTGTTAGCTATGAGTGA	8943
Qy	3783	GAGCATCTGAGAAATTCATGCAACCGTGGAGAAAAGATCAGAGGAGTGTTCGGTGAAG	3842
Dp	8944	GCAACTCAGCAAGTCAATGACAAATGGCCCGCAATATTAACCAAAATTCGTGGACGGG	9003
Qy	3843	AGAGTACTTGTATCCACCCGCTCCTAGGAAAAGCTGTAAAGTGCCACTTATACGATCACT	3902
Dp	9004	AAAAATATGACCTACTCTCCCGTTACGGTATGAGAAATTCCTTGACAGATGTACACCGTC	9063
Qy	3903	TGAAGAGACGCTCTGCGGGGTACATTAACCATGACAGGCCAGGCCCAACGCGATTAAGT	3962
Dp	9064	TGAAAAGAAACAACCCCGGCTACATCATATGACAGGCCGGGACCGCACCGCTATACAT	9123
Qy	3963	CCTATCTGAGGAAACGGTCAGCGGAGTGTACATTAACCACTTCTTGCAAGAACGCA	4022
Dp	9124	CTTATCTGAGAAATCATCAGGAAAAGTTTACGCAACCAACATCCGGGAAAGAACTTA	9183
Qy	4023	CCTACGAATGTAAAGTGGCGCACTACAGCAACAGTATGTGAGCACGGGAAAGAAATGA	4082
Dp	9184	CGTACGATGCAAGTGGCGGCAATTCAGAACCGGAACCTTTACGACCTGTACCGAAATCA	9243
Qy	4083	ACGCGTGAACCTAAGCAAAAACAGTGCATTCGCTTACAGAGCGCAACCAACGAAATGGCTCT	4142
Dp	9244	CGGGCTGACACCGCATCAAGCAGTGCCTGCTATTAAGAGGACCAACGAAGTGGCTCT	9303
Qy	4143	TCAATTCGCGCGGATCTTATTAAGGACACAGAACCATCTAGTGCAGAGTAAATTTGCACATTC	4202
Dp	9304	TCAATTCGCGCGGATCTGATCAGAACACGCGGACACACGCGCCCAAGGAAATTTGACATTTGC	9363
Qy	4203	CATTCCGCTTGAACACGCAAGCTCTCCGGGTTCCGTTAGCTCACACGCGCTACAGTCAACGA	4262
Dp	9364	CTTTACAGCTGATCCCGAGTACTGTGATGTGCTCTGTTGCCACGCGCGGAACCTTAGTAC	9423
Qy	4263	AGTGTTCAAAGGCATCAACCTTCCACTGACTGCATCGCAATCGCAACATTTGCTGACAAACGA	4322
Dp	9424	ACGGCTTAAACACATCAGCGCTCCAAATTAGACAGAACCATCTGACATTTGCTCACACCA	9483
Qy	4323	GAAATTTGGGCTGCGAGACAGCGCAACAGCAAGAAATGATTAACAGGGTCTACATCCAGGA	4382
Dp	9484	GGAGACTGAGGGCAAAACCCGGAAACCAACCTGAATGATCATCGGAAACACGGTTAGAA	9543
Qy	4383	ATTTTTCGTGGGGGAGAAAGGGCTGAGTACGTATGGGGTAAACATTAACCACTCAGAG	4442
Dp	9544	ACTTACCGTGCACCGAATAGGCTCTGGAATACATATGGGGCAATCAGAACAGTAAGGG	9603
Qy	4443	TCTGGGCCAGGAGTTCGACACAGGCGCAACCAATGATGGCCGATAGATCATCATCC	4502
Dp	9604	TCTAAGCCCAAGAGTCTGCAACAGAGAACCTCAGAGATGGCCACAGAAATTAATACACG	9663
Qy	4503	ACTATTATCATCGCATCATCACTGCTCATCTGCTGTGCTGTGCTGTGCTCTTGTCTA	4562
Dp	9664	ATTATATCATCATCGCATCTCTGTGTACACATCTTAGCGTGCATCAGCTCTGTGGGGA	9723
Qy	4563	TCTGGTATGAGCACTGCAATCAGAGCTTGATGCGCAAGGCAAGAAAGAACATCGCTGA	4622
Dp	9724	TGATATATGGGTAACTGTGAGCATTAATGTGCTGTAAAGCGCGGTAGTGGCTGA	9783
Qy	4623	CGCCATACGCGCTTGACCGGAACGCAACGATACCAACAGCACTTAAGCGGTTTGTGCTGCA	4682
Dp	9784	CGCCATATGCGCTGGCCCAATTCGGCGATTCCAACTTTCGCTGCACTTTTGTGCTGTG	9843
Qy	4683	TTGGGCCCAACGCTGAAACATTTTGAAGAACTTTGAACCATCTGTGTGTTTAAACAC	4742

Db	9844	TTAGTGGCGCTAAATGCTGAAACATTACCGAGACATGATTTACTATATGGTCGAAACAGCC	9903
QY	4743	AACCGTTTCTCTGGGACACAGTGTGCAATTCCTCTGGACGGCTTGTATTTCTGTCCGCT	4802
Db	9904	AGCGCTTCTTCTGGGTCAGAGCTGTATATCTCTGGCCGCTGTCTGTCTTATATGGCT	9963
QY	4803	GCTTTTCATGTGTGATGCTTTTATTTATGGTGTGACGGCTGTGCTGTGGGAAGTATAGC	4862
Db	9964	GTTGCTCATGTGTGCTCTCTTTTATGTGTGTGGCGGCGCTTACCTGGCCAAAGTATAGC	10022
QY	4863	CCTTGGAACATGGGACACACTGTGCAAAATGTCCGGGATCCCGTATTAAGCGGTGGTCG	4922
Db	10024	CCTTACGAACATGGGACCACTGTGTCAAATGTGCCAAGATGCCATATTAAGCACTTGTTG	1008
QY	4923	AACGCGAGGTTTACGCCCACTTAACCTGGAGATCAAGTGTCTCATTCGGAATTAAC	4982
Db	10084	AAAGGCGAGGGTACGCCCGCTCAATTTGGAATTAATGTCAATGTCTCTGGAGTTTTCG	10144
QY	4983	CTTCAACTAACAGAGTACGTGACTGTCAAAATTTCAACACAGTCATTCCTTACACCAAG	5042
Db	10144	CTTCCACCAACCAAGAGTAACTTAACGTCAAAATTTCAACACTGTGTGTCCTCCCTCAAG	1020
QY	5043	TTAAATGCTGGGGTCCCTCGAAGTCAAGGCAATCTCAAGGGGATTAACATGCCCGC	5102
Db	10204	TCAATGCTGGGCTCTTGTGAATGTACGCCGCGCTCAGGAGATTAATCTGTCAAGG	10266
QY	5103	TTTTTGGCGGTGTACCTTTTCATGTGGGGAGGCGCACATCTTCTGTGACAGTGA	5162
Db	10264	TCTTTGGAGGGGTGTACCCCTTCATGTGGGAGAGACACATGTTTTTGGGACAGTGA	10322
QY	5163	ACACACAATGAGTGAAGGCGTACGTGAGTTTCGCTCCAGCTGCACTATATGATCACGAG	5222
Db	10324	ACACCCAGATGAGTGAAGGCGGTACGTGAATTTGTCAATGATTTGCGGACATGACCA	10388
QY	5223	TCCGACCTAAAGTTCAACACAGCTGTCTGAAATGTCCGCTCTGTATATGTATAGGCAAC	5282
Db	10384	AGGGGATTTAAGTGCAATACCTCCGCGATGAAGTACGTGCTATATGTGTACGGGAA	10444
QY	5283	CCACCGCGACCTGTGATACGTTTGTCAATGGCGTCAAGCCAGGTTCCCTACGGGACCTGA	5342
Db	10444	CTACCACTTTCCTTATGTGTGTACGTGAAGGAGTCAACCAAGAACTCTTAAGACCTGA	10500
QY	5343	AGGTCATAGCAGGCGCGATATCAAGCCGCTTTTCAACCTTGTACATTAAGTGTGATCA	5402
Db	10504	AAATCATAGCTGTGACCAATTTACAGCATTTGTTTACACATTTGCATCAAGGTGTTATCA	10566
QY	5403	GAAAGGGGCTTGTTTACATCACTACGATTCCTCTGATGTGAGCTATGAACCAAGAGCGT	5462
Db	10564	ATCCGGCGCTGTGTGTACATATGACTTTTCCGGAATTAAGGAGGATGAACCAAGAGCGT	10622
QY	5463	TCCGCGATTTTCAAGCATCTCTCGCTGTATGCTTACAGCATATGATAGGCCGCGACATAC	5522
Db	10624	TTGGAGCATTTCAAGTACTCTCTTGACTAGCAAAAGCTCATGTGCGACGACAGACATTA	10688
QY	5523	GGCTGTGAAGCCCTTCTGTCAAGAACTCCACGTCCTCTTACACCAAGCAAGTATCAGGT	5582
Db	10684	GGCTACCTCAAGCCTTCGCGCAAGAGGTGATTCCTGTATCAAGGAGCGCATCTGAT	10744
QY	5583	ATGAATGTGAGAGAACATCTACAGAGCACTCCGTGACGAAGAACAGCACATTTGTATGT	5642
Db	10744	TCCGAGTGTGGAATAACATCTCAGGCGCGCCACATGTGACGGAACCGGCCCTTTTGGGTGC	10800
QY	5643	AAATTTGAAGTGAAGCCCTGTGGAGCGCTTACATGTGTTATCGGGGACATCCCTATCTCA	5702
Db	10804	AGATTGACATCAATCCGCTTCCAGCGGTGAGCTGTATACGGGAACATTCCTCATTTCTA	10866
QY	5703	TTGACATCCCTGTATGCACTTTTGTGAGATCATCAGATCAACCAAAATTTTGAAGTTA	5762
Db	10864	TTGACATCCCGAAGCGCTTTTATACAGACATCAATGACCACTGTGTCTCAACAGTCA	10922
QY	5763	GCTGCACATGACAGACTGCAATTTATTTGTGACACTTTGGTGTCTCTTAACATTAACGT	5822
Db	10924	AAATGTATGTCAAGTGTGACATTTATTAAGCGGACTTCGAGGAGATGAGCTTACCTTGCA	10988

QY 5823 ACAAGCTGACAGGAGGAGCATTTGTCAGTTCACTCCCACTCCAGCAGCTGTTTGA 5882
DB 10984 ATGATATCCGACCGGAGAGCAATGCTTGTACATTTGCAATTCAGACACAAACCTCC 11043
QY 5883 AGAAGGACACCAACATGTCAGTCCGAGGACAGCATTAACATTAATTTAGCATCGA 5942
DB 11044 AAGAGTGCAGAGTTCATGTCCTGGAGAAAGAGGAGGAGACAGTACATTCAGCAGCGGA 11103
QY 5943 GCCCACAAGCAATTTATTAATGTTTCTGATGCGGAGAGAAATCCACTGTCATGCTGAAT 6002
DB 11104 GCCCAGAGGAGAACTTATGATGCTGCTGTGTGAAGAACAACTGCAATCAGAAAT 11163
QY 6003 GTAAACACCGGCGGACACATTAATGAGAACACAACTAAAGTGACCAAGATTTCCAG 6062
DB 11164 GCAACACACAGCTGATCATGTCGAGACACCCGACAAATAAGACAAATTTCCAG 11223
QY 6063 CGGACGTTTCCAAACATCTTGGAACTGCTGCTTGCATGCTTTGGGGAGCATCATCC 6122
DB 11224 CCGCATCTCAAAAACCTTCATGAGAGTTGGCTGTTGCCCTTTTGGGGGCGCTCGTCC 11283
QY 6123 TCATTTGTTAGACTTATTAATGTTGTTGCTGACGCTCATGCTTATAACACAGTATGAT 6182
DB 11284 TATTAAATTAATGAGCTTATGATTTTGTCTGACAGATGCTGATAGCAGAAAGAT 11343
QY 6183 GACTGAGCGGAGCACTGA-CATAGCGGTAAACCTGATCTTCCGAGGAGCGTGT 6241
DB 11344 GACCGCTACGCCCCCAATGACCCGACACGAAACCTCGATGATCTCGAGGAGCTGAT 11403
QY 6242 GCATATGTCACAGCGCGCT-----TGACACT 6268
DB 11404 GCATATGTCACAGCGCTGATTAATGATCCCGCTTACCGGGGACATATAGCAACCC 11463
QY 6269 AAAAAGTGAATTTTCCGAGGAGACAGTCAATGCTGTGCAATGT 6318
DB 11464 AAAAAGTGAATTTTCCGAGGAGACAGTCAATGCTGTGCAATGT 11513

RESULT 6
AR096557
LOCUS AR096557 11663 bp DNA linear PAT 08-SEP-2000
DEFINITION Sequence 7 from patent US 6008035.
ACCESSION AR096557
VERSION AR096557.1 GI:10025467
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 11663)
AUTHORS Johnson,R.E., Davis,N.L. and Simpson,D.A.
TITLE System for the in vivo delivery and expression of heterologous
JOURNAL genes in the bone marrow
FEATURES Patent: US 6008035-A 7 28-DEC-1999;
source Location/Qualifiers
1..11663
/organism="Unknown"
BASE COUNT 3289 a 3105 c 2910 g 2359 t
ORIGIN

Query Match 37.0%; Score 2397.6; DB 6; Length 11663;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3752; Conservative 0; Mismatches 1984; Indels 74; Gaps 9;

DB 5827 AAAAATCTACGCCCCGCTGTCAGACGTCGAAAGAGAAAGCTCAAACTCAGTACC 5886
QY 760 AATTATCGGCTCTGAAGAAATTAAGAGAGTATCAATCAGAAAGTGAATAATGA 819
DB 5887 AGATGATGCCACCGAAGCCAAAGCAGAGTACGATTCGAAAGTGAAGAAACAGA 5946
QY 820 AAGCAATTAAGAGGAGCGACATATTTCTGATTTGGACATATCTATCAAGAACTGA 879
DB 5947 AAGCATTAACCATGAGCGACGCTGTTTCAAGGCTACGCGTGTATTACTGCCACAG--- 6003
QY 880 ATCTGTGAGTGTATTAAGATCAATTAATCTGTGATCAATCTGTAACGCTAAATTA 939
DB 6004 ATCAGCAGAAATGCTATTAAGATCACTTACCCGAAACCATGTATTTCCAGAGTGTACAG 6063
QY 940 ACAGCTTAATCTGCGAGGTGCGGCTTAAACGTGCAACTTAATGATCCAGAAAT 999
DB 6064 CGAATCTCTGACCCAAAGTTGCTGTAGCTGTTTGTAACTATCTGATGAGAAAT 6123
QY 1000 ACCCTACAGTACGACGATTTATTAACAGATGAATAGATGCGTATCTTGACATGCTGG 1059
DB 6124 ACCGACGTTAGATCTTATCAGATCCGACAGATGCAATGCTTACTTGGATATGCTAG 6183
QY 1060 ACGGCGATGCTGCTGTAGATACAGCACTTTTGTCCGCTAACTGAGAACTACC 1119
DB 6184 ACGGACAGTCTGCTGCTGTAGATACAGCACTTTTGTCCGCTAACTGAGAACTACC 6243
QY 1120 CAAAGAGCATGATTAATTTGACGACGATTAAGATAGCGCTTCCCTATACAGA 1179
DB 6244 CGAAAGACACAGATTAAGAGCCCAAACTCCGACGTGGCTTCATCAGCGATGAGA 6303
QY 1180 ATACATTACAAAATGATTTGGCTGACGCTTAAAGAAATGCAAGTTGCAACCAATGC 1239
DB 6304 ACAGTTGCAAAAAGTCTCTATTCGCGACCTAAAGAAATCTGACAGTCAACAAATGC 6363
QY 1240 GAGAAATTAAGTCTTGTAGATTTGCGGCGCATTTAATGTTGATGTTTCAAGAAATGCAAT 1299
DB 6364 GTGAATCTGCAACACTGATCTGACGCACTTCAACGTTGAATGCTTTCGAAATATCAT 6423
QY 1300 GCATGATGAGTACTGGGATACCTTTGGGATTAACCTTATGGCTATCAACAGAGAAC 1359
DB 6424 GCAATGACGATTAATGGAGAGATTTGCCGAAAGCCAAATTAAGATCACTACGAGTTCC 6483
QY 1360 TTACGCAATTAATGTCACAAAGCTGAAGAGGCGGAAAGCAGACATTTGTCGATATCTC 1419
DB 6484 TTACGCAATTAATGTCACAAAGCTGAAGAGGCGGAAAGCAGACATTTGTCGATATCTC 6543
QY 1420 ATAAATCTAAACCGTTGACAGAGATCAATGATCAATTCGTATGATGATTAAGAGAG 1479
DB 6544 ATAAATTTGTCCTCAATGTCAGAAAGTCTTATGATTAATTCGTATGATGATTAAGAGAG 6603
QY 1480 ATGTCAAAGTATCTCCGCGGACGAAACATACAGAGAGCGGCTTAAGTGCAGTTATTC 1539
DB 6604 ACGTGAAGATTAACCTGTCGACGAAACACAGAGAGAAAGCCGAAAGTCAAGTGTATAC 6663
QY 1540 AGGCTGACATCCCTGCTGTAACGCTTACCTTTGCGGATTCATTCGGAATTAAGTCCGTA 1599
DB 6664 AAGCGGAGAACCCCTGTCGACGACGCTTACTATTCGCGGATTCACCGGAGATTAAGTCCG 6723
QY 1600 GACTGAATGCGGTCTTCTGCAAAATATCAATCTCTTTCGACATGTCAGCGGAGATT 1659
DB 6724 GCGTTACAGCGGTTTCTCTACCAACATTCACAGCTCTTTCGACATGTCGCGGAGAGACT 6783
QY 1660 TTGATGCGATTAATTTGCTGAACATTTTCAACAGCGGACCCAGTATTTGGAACCGAACATTC 1719
DB 6784 TTGATGCGATTAATTTGCTGAACATTTTCAACAGCGGACCCAGTATTTGGAACCGAACATTC 6843
QY 1720 CGTCTGTTGATTAAGAGGAGAGACGACCTATTCGCTTTCGCGCTGATGATCTTTGAGG 1779
DB 6844 CTTCTGTTGACAAAGCAGAGACGACCTATTCGCTTTCGCGCTGATGATCTTTGAGG 6903
QY 1780 ACTTAGGTGTGACCAACCGCTCTTGAATTTGATTAAGAGGCGGCTTTCGCAATATCAAT 1839
DB 6904 ACTTAGGTGTGACCAACCGCTCTTGAATTTGATTAAGAGGCGGCTTTCGCAATATCAAT 6963

QY 1840 CTGTGACCTTACCTACAGAAAGAGTTTAAATTTGTGTCATGTAATCCGGTATGT 1899
DB 6964 CCACCATCTGCGCCACGGGTATCCGTTTCAATTCGGGGCGATGATGAATCCGAAATGT 7023
QY 1900 TCTTAACTGTTTGTCAACACACTATGATATATGATTTGTAGCAGAGTACTACGTG 1959
DB 7024 TCTTCACTCTTTTGTCAACACAGTTTGAATGTGTTATCCGACAGAGTATTTGAGG 7083
QY 1960 AACGTTTAAACAGCTGACGCTGCGGCGCTCTATGCGGACATTAACATAGTCATGTG 2019
DB 7084 AGCGCTTAAACGTCCAAATGTGACATTTATGCGGACGACAACTTATATACGAGG 7143
QY 2020 TCGTCTCCGACACTTGTATGCGGAGAGATGCGCACTTGTGACATGAAATGAAAAA 2079
DB 7144 TAGTATCTGACAAATAATGCTGAGAGGTGTGCGACCTGCTCAACATGAGAGTTAAGA 7203
QY 2080 TTATGATGACGTTATTTGATCAAAAGCACCTTACTTGTGGGGATTTATCTGTGTG 2139
DB 7204 TCATTGACGACGTATCGGCGAGACACTTACTTCTGCGGTGATTCATCTTGCAAG 7263
QY 2140 ACCGATTAACAGGCGACGCTGACAGTGTGACACCTCTTAAAGGCTTTTAAAGCTTG 2199
DB 7264 ATTGCTTACCTCCACAGCGGTGTGCGTGTGCGGACCTCTTGAAGGCTGTTTAAAGTTGG 7323
QY 2200 GAAAAACATTGCGCACTGATGATACCAAGACTGCGCGCGCGGACATGATGATG 2259
DB 7324 GTAAACCGCTCCGACCGACGATGAGCAAGACGAAAGACGAGCGCTCTGTCTAGATG 7383
QY 2260 AAGCAATGCGATGAAACAGAAATTTGAAATTAACGAGAGTATGAAAGCGGTAGATCA 2319
DB 7384 AAACAAAGCGGTGTTTAAAGTATTAACAGACCTTATGACATGTGGCGGTGCAATC 7443
QY 2320 GATACGATCATATCTGCGAGCGCTGATCATCATCTCTGTCCAGTTAGCCGAAAGCG 2379
DB 7444 GGTATGAGTATGACAAATCAACACCTGTCTGTGCAATGAACTTTTGGCCGAGACA 7503
QY 2380 TTAAGAACCTTCAAGACATTAAGAGGAGCGCCAACTCTTACGCGCTGACCTTAATAG 2439
DB 7504 AAAGAGCACTTTCAGACCATCAAGAGGGAATTAAGCATCTCTACGAGTGTCTTAATAGT 7563
QY 2440 TGACGTATGAG-----CAGCACTTACCCACCGCGAGATGTTTCCA 2482
DB 7564 CAGCATATGATCTTCACTGATCACTAATACCAACACCACTGATGAGATTCCT 7623
QY 2483 TACCTCAGCTGAATCTTTCACCACTTAACCTTAATCGATGCTTACCGAGATCCA 2542
DB 7624 TTAACATGCTGCGCGCGCGCTTCCAGCGCCCACTGCACTGTGAGGCGCGAGAA 7683
QY 2543 AACCTCTAGGCGCGCTGAGGCGGTTTGGCGCCCGCTGCTCAATGGAAT 2602
DB 7684 GAGGACAGGCGCGCGGATGCTGCGCGCATGCGCTGCTTCCAAATTCGACAACTGA 7743
QY 2603 CTTAGAGCTGATATGATCACTTGAATTTCAAAAC--GATCAGCTTATCCGCGCGAG 2660
DB 7744 CCACAGCGGTAGTGTCTTATGATGACAGCAACTAGACTTAACCCACGCCAC 7803
QY 2661 GTCCACCGCCCAAGAAAGAAAGATGCTCTTAACCAAACTTACTAGCTTAAAGAA 2720
DB 7804 GCGCGCGCGCGCGAGAAAGAGCGCGCCAAACCAACCGAAGCGAAGAAACCA 7863
QY 2721 AGAAGCAGCAAGCCAAAGAGACGAAC--GCAAGCTTAAACAGAGAAAGCAACGTA 2777
DB 7864 AAACACAGAGAAAGAAAGAAAGCACTGCAAAACCAAAACCGGAAAGAGACGCTGA 7923
QY 2778 TGTGTATGAGTTGAGTGTGCAACAGATTTCCGATC--ATGCTGAACGCGCAAGTGA 2834
DB 7924 TGGCACTTAAGTTGAGGCGGACAGACTGTTGAGTCAAAATATAGAGCGGAGATGCA 7983
QY 2835 ATGATATGCTGCTGTGTGAGAGAGGCTGATAAACACTCCAAGTTGAGGAAAA 2894
DB 7984 TCGGCGACGCACTGCGCATGAGAAAGAAAGTAAATCACTCCACGTGAAGAAAGACTA 8043

QY 2895 TTGATATGACAAATTACGCGCGGTGAATTTGAAGAAAGCTAGACTATGCACTTGGAGT 2954
DB 8044 TTGACACCCGTGTGTATCAAAAGCTCAAAATTCACCAATCTCTACGATACGACATGAGT 8103
QY 2955 ACGCGACGTTTCCCGAAACATGAATCAGACAGCTGCACTACACGAGCAAAACAC 3014
DB 8104 TCGCAGACTTCCCGTCAACATGAGAGTGAAGGCGTTTCACTTACACAGTAAACCTTG 8163
QY 3015 CGGCTTCTAACAATGAGACCAAGCGCGGATGCTGAGTATGAGAAATGGGATTTTACCGTAC 3074
DB 8164 AAGGTTTCAACCTGACCAAGAGGCGGTGAGTATGAGAGGAGATTTTACATTC 8223
QY 3075 CGAGAGAGTGGGCGGAAAGGCGACAGCGGAAAGCCATCTGTGACAAACAGAGCAGAG 3134
DB 8224 CCGCGGAGTATGAGAGGAGAGAGAGTGTGTGCTGCTGATATGATTAATCTACAGCGCGG 8283
QY 3135 TTGTGCTATTTGTTAGAGGTGCAATGAGGCGCGGTACGCGCTTTTCACTGTCTCA 3194
DB 8284 TTGTGCGATAGTCTTCCGAGGGGCTGATGAGGGAACAAAGCCGCTTTCGCTGTCA 8343
QY 3195 CTTGGAACCAAGAGGGGTGACATTTAGGATTAACCCCGAAGTTTCAACCGTGT--- 3251
DB 8344 CTTGGAATGCAAGAGGAAACAAATCAAGACAAACCCGGAAGGACAGAAAGTGTGTG 8403
QY 3252 -----CACTAGTTACAGCGCTATGCGCTTTCGAAATGTCACGTTCCCATGCGCAAA 3305
DB 8404 CTGACACCTGTGTACGCGCATGTGTCTTGTGAAAGTGAAGTCTTCCATGTCAATCCGC 8463
QY 3306 CACCGTGTCTATTCACTGACGCGAGAACGAAACCTGACGTGTGAGAGAAAGCTGCG 3365
DB 8464 CGCCACATGCTTACCGCGGAAACCATCAAGAGCTGTGATCATCTCGAAGAAAGCTGA 8523
QY 3366 ACAATCAATTTAGACACGCTGCTGAGAAAGCTTTGAAT--GTCTATACGCGCGC 3422
DB 8524 ACCAGAGCGCTTACGACACCTGTCTCAAGCATATGCGGTGCGATGCTGCGCAAA 8583
QY 3423 CCAAGCAAGGATTAACGATGATCACTGACAGCGCCCTTACCTGGGGTTTCTGCGCGT 3482
DB 8584 GTAAAGAGGCTCACTGACATCTTATCTTGAACAGCGCTTACTTGGGACATGTCTGT 8643
QY 3483 ATTCAGACACTCAACGCGGTGTTTCAAGCCCAATTAATTAATGAGAACGTGTGGAAGT 3542
DB 8644 ACTGTACCATATGAAACCGTGTGTTAGCCGATTAATGAGAGAGTGTGGAAGTGAAG 8703
QY 3543 CTGATGATGATCGATTAATGATCCAGGCTGTGCGCAATTTGGGCTTCAATACAGAGCA 3602
DB 8704 CGGACGACAAACATAGCATAGACATCTTCCGCGCATTTGATGATGACAAAGCGGAG 8763
QY 3603 CTGCGGATGTACCAAAATTCGTTTACATGTCTTTCGACACAGACCATGACATCAAGAG 3662
DB 8764 CAGCACTCAAAATTAATGATCCGCTTACATGTGTCTGACAGAGATCATCTGTCAAGAG 8823
QY 3663 ACAGATGAGAAAAATAGCTATGACACATCTGACACCTTGCCTTGTGCGCCCAAG 3722
DB 8824 GCACATGATGATGATCAATGATGACAGACCTCAGAGCGGTGTAAGAGCTTATGCTCAAG 8883
QY 3723 GGTACTTCTTATGACTCAATGTCTCCAGGTGACAGTGTAAACGCTCATGATACAGAGC 3782
DB 8884 GATACTTCTCTCGGAAAGTGTCTCCAGGGGACAGGTATACGTTAGCATAGAGATA 8943
QY 3783 GAGCATGTGAATTCATGACCGGTGAGAAAGAAAGATCAGAGAGAAAGTTTGTGGTAGAG 3842
DB 8944 GCAACTCAGCAAGCTCATGCAATGCGCGCCAAAGATTAACCAAAATTCGTGGACGCG 9003
QY 3843 AGGATACTTGTTCACACCGCTCATGAGAAAGCTGTAAAGTGCACCGTTTACATCACT 3902
DB 9004 AAAAATATGACCTTACCTCCGTTACCGTAAAGAAATCTTTCAGAGTATACGCGTTC 9063
QY 3903 TGAAGAGAGCTCTGCGGGTATCAATTAACATGACAGGCGCGCACCGCTTATAGT 3962
DB 9064 TGAAGAAACAAACCGCGGTATCACTTATGACAGGCGCGGACCGCATATCAT 9123
QY 3963 CCTATCTGAGAGAGCGTACAGGAGAGTATTAACCACTTTCGCAAGAAAGCTCA 4022

Db 9124 CCTATCGAGAAATCATCAGGAAAGTTTACGGGAACCCATCCGGGAAGAACATTA 9183
Qy 4223 CCTACGAATGTAATGTGGGCACTACAGACAGATATGTAGACACGGAAAGATGA 4082
Db 9184 CGTACGAGTGAAGTGGCGCATTTACAGACCGGAACCGTTACACCCGTGCCAAATCA 9243
Qy 4083 ACGGCTGCACTAAAGCAAAACAGTGCATTTGCTTCAAGAGGACCAAAAGGAGGTCT 4142
Db 9244 CGGGCTGACCGCCATCAGAGATGCTGCTTATAGAGCGACCAAGAGTGGTCT 9303
Qy 4143 TCAACTGCGCGGATCTTATTAAGCACACAGACCACTAGTCAAGTAAATTTGCAATTC 4202
Db 9304 TCAACTGCGCGGATCTGATCAGACACCGCGACCAACGGCCCAAGGAAATTTGATTTGC 9363
Qy 4203 CATTCGCTTACACCGACAGTCTGCCGGTTCCTTACCTACACGCTTACAGTACGA 4262
Db 9364 CTTCACAGCTGATCCCGAGTACTGCAATGCTCCCTGTTCACACGCGCCGAACGTATAC 9423
Qy 4263 AGTGTTCAAAGGATCACCCTCCACCTGACCTGCAATGCAACCAATTTGCTGACAA 4322
Db 9424 ACGGCTTTAAACATCAGCTCCTTCAATTAGACAGACCATCTGACATTTGCTTACACCA 9483
Qy 4323 GAAATTTGGGCTGCGAGCAGACCAACAGAGATGATTAAGGCTTACATCCAGA 4382
Db 9484 GGAAGCTAGGGGCAAAACCCGAAACCAACCACTGAATGATCATGGAAACAGGTTAGAA 9543
Qy 4383 ATTTTCTGTGGGCGAGAAAGGCTGAGTACGTATGGGGTAAACATTAACCACTAGAG 4442
Db 9544 ACTTCACGCTGACCGAGATGGCTGTGAATCATATGGGGCAATCAACAAACAGTAAAGG 9603
Qy 4443 TCTGGGCGCAGAGTGGGCAACGAGCCGACCAACATGATGGCGGATGATCATCTCC 4502
Db 9604 TCTATGCCAGAGTCTTGACACAGAGACCTTCAGAGATGGCCACAGAAATGATGACGC 9663
Qy 4503 ACTATATCATCGGATCCAGTCTACATCTGATTTGCTGTGTGTCTGCTCTTGTGTA 4562
Db 9664 ATTAATATCATCGGATCCGTGTATACATCATTTAGCCGTGACATGCTGCTGTGGGA 9723
Qy 4563 TCTGTGAGGCACTGATCATCAGACGCTTGCATGCGCAAGAGAGATGCTCTGA 4622
Db 9724 TGAATATTTGGGTAATCTGTTCAGATTAATGCTGTAAAGCGCGGTAGTGGCTCGA 9783
Qy 4623 CGCCATACGGCTTGCACCGAAGCAAGGTAACCAACAGATTAAGCGGTTTGTGCTGA 4682
Db 9784 CGCCATATGCTGCGCCCAATGCGGTATTCACCTTCCGTGCACTTTTGTGCTGTG 9843
Qy 4683 TTCGGCAACCAACGCTGAAACATTTGGAGAAACCTTGAACATCTGTGTTTAAACA 4742
Db 9844 TTAGGTGGGCTAAATGCTGAACATTCACCGAGACATGATTAATGATGCAACAGCC 9903
Qy 4743 AACCGTTTCTGTGGGACAGTTGTGATCTCTGTGACGCGCTTGTATTTCTGTCCGCT 4802
Db 9904 AGCGGTCTTCTGTGGGCTAGCTGTGTATACCTGTGCGCGCTGTCTTCTTAATGCGCT 9963
Qy 4803 GCTTTTCAATGCTGATGCTTTTATATGTTTGAAGGCGTCTGCTGGGGAAGTGAAG 4862
Db 9964 GTTGTCTATGCTGCTGCTTTTATAGTGTGCGGCGCTTACTGTGCGAAAGTGAAG 10023
Qy 4863 CTTTGAACATGCAACACTGTGCCAAATGTTTCCGGGATCCCGTAAAGCGTGTGTCG 4922
Db 10024 CTTGAAGACATGCGACACTGTTCCAAATGTGCCACAGATACCGTATAGGCACTGTGTG 10083
Qy 4923 AACGCGAGTTTGAAGCGCACTTAACCTGTAGATACGCGTGTCTCATGGAATTAAC 4982
Db 10084 AAAAGGCGAGGTAGCGCCCGCTCAATTTGGAGATTAATCTGTATGCTTCGGAAGGTTTGC 10143
Qy 4983 CTTGAATTAAGAGAGTGAAGCTGCAATTCACAGATCATCTTCCACACAG 5042
Db 10144 CTTTCAACCAACAGATGATTAACCTGCAAAATTAACCACTGTGTGCTTCCCTTAAAG 10203
Qy 5043 TTAATGCTGGGCTCTCGAGTGAAGGATCCTCAAGGCGATTAACATGCGCGG 5102

Db 10204 TCAATGCTGCGGCTCCTTGAATATGTCAGCCCGCTCAGAGATATACCTGCAAG 10263
Qy 5103 TTTTGGCGGATGCTATACCTTTTCAATGTTGGAGGCGCAATGCTTGTGACGTAGA 5162
Db 10264 TCTTTGAGGAGGTATACCTCTTCAATGTGGAGAGAGCAATGTTTGTGACGTAGA 10323
Qy 5163 ACACACAATGAGTGAAGGCTTACAGTGTGCTCCAGACTGACATATAGTACAGCAG 5222
Db 10324 ACAGCCAGATGAGTGAAGGCTTACAGTGTGATTTGATAGTATGTCGACACAGCGC 10383
Qy 5223 TCGCATTAAGTTTCAACAGCTGCTGAAAGTGGCTGCTGTATATGATAGCGCAACA 5282
Db 10384 AGGCGATTAAGTGTGACATCTGCGGATGAAGATGAGATGCGTATGTGTATCGGGA 10443
Qy 5283 CCACCGGCAACCTGGAATGCTTGTCAATGGCGTACGCGAGGTTCTCAAGGACCTGA 5342
Db 10444 CTACCAATTTCTAGATGTATGATGTAAGAGAGTACACAGGAACCTTAAAGACCTGA 10503
Qy 5343 AGTCAATGAGAGGCGCATATCAAGCCGCTTTTCAACCTTTTGAACATTAAGTCTCATCA 5402
Db 10504 AAGTCAATGCTGAGCAATTTTACAGCATTTTATACCATTCGATTCACAGGTGTATCA 10563
Qy 5403 GAAAGGCGCTTGTATTAACATGACTTCCCTGATATGAGCTATGAACAGAGCGT 5462
Db 10564 ATCGCGGCTGTGTATCAACTATGATCTTCCGGAATAGGAGCGATGAACAGAGCGT 10623
Qy 5463 TCGCGATTTTCAAGCATCTCTGCTTGAATGCTACAGATAGTACCGGACTGACATAC 5522
Db 10624 TTGAGACATTTCAAGCTACTCTTGAATGACAAAGCTCTATGCGACAGACATTA 10683
Qy 5523 GGCTGTGAAGCCTTCTGTCAAGAAATCCACGCTCCCTTACACCAAGCATTAAGGCT 5582
Db 10684 GGTACTCAAGCTTCCGCGCAAGAGCGTATGCTCCGTACAGSAGCGCGCATCTGGAT 10743
Qy 5583 ATGAATGTGAAGAACATCTCAGAGCAGCCCTGCAAGAAACAGCAATTTGATGTA 5642
Db 10744 TCGAGATGTGAAGAAACATCTCAGGCGCCGACAGGAAACCGCCCTTTTGGGTGCA 10803
Qy 5643 AAATGAAGTGAAGCCTCTGCGAGCGCTTAACTGTGCTTACGGGCAATCCCTATCTGA 5702
Db 10804 AGATGCAATCAACCTGCTGAGCGGTGACCTGTATACGGGAACATTTCTTA 10863
Qy 5703 TTGACATCCCTGATGACAGCTTTTGTAGATCATCAGATCAACCAATTTTGAAGTGA 5762
Db 10864 TTGACATCCCGAAGCGCTTATATCAGACATGAGTGAACCACTGCTCAACATCA 10923
Qy 5763 GCTGCAAGTGAAGACTGATTAATTTCTGCAACTTTGTGCTTCTTAACTTAAGT 5822
Db 10924 AATGTATGTAGTGAAGTGAATTAATCAGCGGCTTGGAGGATGCTTACCTGTGAGT 10983
Qy 5823 ACAAGCTGACAGGAGAGGCAATTTGTCAGTTCACTCCCACTCCAGACAGCTGTTTGA 5882
Db 10984 ATGTATCGAACCGGAAGCAATGCTGTGACATTTGCAATTTGAGACAGCAACCTCC 11043
Qy 5883 AGAAGCGACACACATGATGCTGCTGAGAGCATTAACATTAATTTTGAACATGCA 5942
Db 11044 AAGAGTGAAGATTCATGCTCTGAGAAAGAGAGCGTGAAGTGAACCTTCAAGCAG 11103
Qy 5943 GCCCAAGCAATTTTATATGTTTCTGATGCGGCAAGAGTCACTGCAATGCTGAAT 6002
Db 11104 GCCCAGAGGAACTTATTTATGCTGTGTGTGTAAGAGACAACTGAATGCAAAAT 11163
Qy 6003 GTAACCAAGCGCGGACACATTAATGGAACACATTAAGTGAACCAAGAAATTCAGG 6062
Db 11164 GCAACCAAGCTGATCAATGTGAGCAACCCGCAACAAATGACCAAGAAATTCAG 11223
Qy 6063 CGGCAATTTCAAAACATTTTGAACCTGCTGCTTGAACCTTTTGGGAGACATATCCC 6122
Db 11224 CCGCATCTCAAAAACCTTCAATGAGTGTGCTGTTTGGCTTTTGGGCGCGCTGCTGCG 11283
Qy 6123 TCAATGTTGAGCACTTAATGATGTTGTGCAAGCTTATGCTTATTAACACAGTATAG 6182
Db 11284 TATTAATTAAGGACTTATGATTTTGTCTGACAGATGATGCTGACAGACAGAGAT 11343

QY	6193	GACGTAGCGGGAGCACTGA-CATAGCGGGTAAACTGANTGTACTTCCGAGGAAGGGTGT	621
Db	11344	GACCGCTAGCGCCCAATGACCACCGACCAACTGATGTCTTCGAGGAACCTGATGT	11404
QY	6242	GCATATGCGACGCGCCGT-----TGACACT	6266
Db	11404	GCATATGATGATCAGCGCTGGTATATATAGATCCCCGCTTACCCGGGCAATATAGCAACAC	11464
QY	6269	AAACTCCGATGTATTTCCGAGAACACAGTCGATTAATGCTGTGACGT	6318
Db	11464	AAAACTCAGACGTAATTTCCGAGAACGCGATGCAATATGCTGCGACGT	11513
RESULT 7	BD071122	11663 bp DNA linear	PAT 27-AUG-2002
LOCUS	BD071122		
DEFINITION	System for the in vivo delivery and expression of heterologous genes in the bone marrow.		
ACCESSION	BD071122		
VERSION	BD071122.1	GI:22616725	
KEYWORDS	JP 2001515348-A/3.		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 11663)		
AUTHORS	Johnson, R.E., Davis, N.L. and Simpson, D.A.		
TITLE	System for the in vivo delivery and expression of heterologous genes in the bone marrow		
JOURNAL	Patent: JP 2001515348-A 3 18-SEP-2001; UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL OS Girdwood S.A.		
COMMENT	PN 18-SEP-2001		
	PD 18-SEP-2001		
	PF 18-FEB-1998 JP 1998536760		
	PR 19-FEB-1997 US 08/801263		
	PI ROBERT E JOHNSTON, NANCY L DAVIS, DENNIS A SIMPSON PC		
	CI12N15/86, CI12N15/33, CI2N7/01, CI2N5/10, A61K39/12, A61K48/00 CC		
	System for the in vivo delivery and expression of heterologous CC genes in		
	CC the bone marrow		
	FH Key	Location/Qualifiers	
	FT source	1. 11663	
		/organism="Girdwood S.A."	
FEATURES	source	1. 11663	
		/organism="unidentified"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:32644"	
BASE COUNT	3289 a 3105 c 2910 g 2359 t		
ORIGIN			
Query Match	37.0%; Score 2397.6; DB 6; Length 11663;		
Best Local Similarity	64.6%; Pred. No. 0;		
Matches 3752; Conservative	0; Mismatches 1984; Indels 74; Gaps 9;		
QY	580	ACTGACGGTATGAAACGGGAGCGTATATTTCTCATCGGAAACAGGCCAAGTACCTTC	639
Db	5707	ACTGCTAACCGGGGTAGGTGGTACATATTTTCAGCAGACACAGCCCTGGCACTTGC	5766
QY	640	AACAGAAATCAGTACGTCAATGTAACTACAAACCTATATGTGATCGGGCCGTCCATG	699
Db	5767	AAAAGAAATCCGTTCTCGACAAACGCTTACGAACCGACCTTGGAGCGCAATGTTCTGG	5826
QY	700	AGAAGTATTAAGCCCCCGCGCTCGATCTCGAAGAGAAATGTTACAGAAACTGC	759
Db	5827	AAAGATCTAGCCCCCGCGGTGCTCGACACGTGAAAGAGAAACAGCTCAAACTCAGTACC	5886
QY	760	AATTATGCGCTCTGTAAGAAATAGAACAGGTATCATCAAGAAAGTAAATATGA	819
Db	5887	AGATGATGCCCAACGAAACCAAAACAGGTACCACTCTCGAAATAGTAAACCGA	5946
QY	820	AAGCAATTACAGCGGAGCACTCATTTCTGATGTGGGACATATCTATCATCAGAGTGA	879

D	5947	AAGCCATTAACCACTGAGGACTGCTTTAGGGCTACGGCTGTATTAACCTCTGACAG--	6003
Q	880	ATCTGTGAGTGTTCAGAGTCAATTATCCTGTACCAATCTACTCGTCAACGGTAATTA	939
D	6004	ATCAGCCAGAAATGTCTATATAGATCACTACCCGAAACCATGTGATTTCCAGCAGTGTACAG	6065
Q	940	ACAGTTTACATCTGCAGAGGTCCGGGTAAACGTGCACCTTACTGTTATCCAGAGAAATT	999
D	6064	CGAATCTACTGACCCAAAGTTTGGTGTAGCTGTTTGTATACACTATCTGCATGAGAAATT	6122
Q	1000	ACCCACAGTACGCCAGTATTTGTATTAACAGATGATATACAGCGGATCTTGACATGGTG	1055
D	6124	ACCCGACGGTAGCATCTTATCAGATCACCGACGAGTACGATGCTTACTTGGAATGTGTAG	6183
Q	1060	ACGGCGATCGTGCTGTCTAGATACAGCCACTTTTGTCCGGCTAAACTGAGAACTACC	1119
D	6184	ACGGGACAGTGGCTTGCCCTATATCTGCACTTTTGGCCCGCCAGCTTATGAAATTAAC	6243
Q	1120	CAGAAAGCATAGCTATTTGACAGCCAGAGATTAAGTACGCCCTCCATCGCTTATACGA	1179
D	6244	CGAAAGACACAGAGATATAGAGCCCAACATCCCAAGTGGCTTCACTACAGATGACGA	6303
Q	1180	ATACATTACAATATGTATTTGGCTCAGCTACTATAAGGAATTTGACGTTACCAATAGC	1239
D	6304	ACAGCTTGCAAAAGTGCTCACTTCGCCGACTTAAAGAACTGTGACCTCACACAAATGC	6363
Q	1240	GAGAAATTCTGTGTAGATTCCGCGCGCATTTAATGTTGATTTGTTTCAAGAAATACCAT	1289
D	6364	GTAAGCTGCCAAGCTGCACTCAGCGACATTCMAAGTTGATGCTTCGAAATATATGAT	6423
Q	1300	GCATGATGAGTACTGGGATACCTTTGCGATAACCTTATTCGCTTAACTACAGAGAACG	1359
D	6424	GCAATGACGATATTTGGAGAGAGTTGGCCGAAAGCCAAATTAGATACATCTAGATTGC	6483
Q	1360	TTAGCGAATATGTGACAAAGCTGAAGGCGGAAAGCAGACGACTGTGTTGCCAATCTC	1419
D	6484	TTACCGCATAGTGGCCAGACTGAAGGCCCTTAAGGCCGCGCACTGTTCGCAAAAGACG	6543
Q	1420	ATTAATCTTAAACCGTTGCAGAGATATCAATGATATCAATTCGTATGATTTAAAGAGAG	1479
D	6544	ATAATTTGCTCCCATCTGCAAGAAAGTGCCTATAGTATGTTGTATGACATGTGAAAGAG	6603
Q	1480	ATGTCAAAATTACTCCCGGACGAAACATACAGAGAGCGGCTTAAGTGCAGTTATTC	1539
D	6604	ACGTGAAGTTACACCTGGCAGCAAGAAACACAGAAAGAACCGAAAGTACATGATATAC	6663
Q	1540	AGGCTGCAGATCCCTTGTCTACGGCTTACCTTTGCGGATTCATCGGGAATTAATCCGTA	1599
D	6664	AAGCCGCGAAGCCCTGGCGACCGCTTACCTATCGGGATATCCACGGGAGATTATGCGCA	6723
Q	1600	GACTGAATGGGAGTCTCTGCGCAATATCCATCTCTTCGACATGTCAAGCGGAAGATT	1659
D	6724	GGCTTACAGCGGTTTGTCTACCAATTCACAGCTCTTTTGACATGTGCGCGGAGACT	6783
Q	1660	TTGATGCGATTAATTTGCTGAACATTTCCACACGCGGACCCAGTATTTGGAACGACATCG	1719
D	6784	TTGATGCAATCATAGCAGAACATCTTCAAGCAAGGTGACCCGTAATCTGAGACGGATATCG	6843
Q	1720	CGTGTGTTGATTAAGGAGAGACGACGCTATCGCATTTTCGGCGTTGATGATCTTTAGAG	1779
D	6844	CCTGTGTCGAAAGCCAGACGACGCTATGCGTTAACCGGCGTATGATCTTTGGAAG	6903
Q	1780	ACTTAGGATGAGCAACCGCTTTAATTTGATATGAGAGGGCGGCTTGCGGAATATTCAT	1839
D	6904	ACCTGGATGTGACCAACCACTACTCGACTTGATCGAGTGGCGCTTTGGAGAAATATCAT	6963
Q	1840	CTGTGCACTTAACCTACAGAAACGAGTTTAAATTTGGTGCATGATGAATCCGGTATGT	1889
D	6964	CCACCCATCTGCGACCGGTAACCGTTTCAAATTCGGGGGAGATGATGAATCCGGAAATGT	7023
Q	1900	TCTTAAGCGTGTGTGTCAACACTAGTCAATATCATGATTTGCTAGCAGAGTACTACGTG	1959

Dh 7024 TCCTCAGCCTCTTGTCAACACAGATTCTGAATGTCTTATTCGCCAGACAGATTGGAGG 7083
Qy 1960 AAGCGTTAACCAAGCTCAGCGTGGCGGCTCTATTCGGGAGAGATAAGTATGTGATGTG 2019
Db 7084 AGCGGCTTTAAACCTTCAAATAGTACAGATTATTCGGGAGAGCAACATTATACAGAG 7143
Qy 2020 TCGTCTCCGACCTTGTATGGCGGAGATGCGCCACTTGGCTGACATGGAGATAAATA 2079
Db 7144 TAGATCTGACAAAGAAATAGTCTGAGAGTGTGCACTGGCTCAACATGGAGTTAAGA 7203
Qy 2080 TTATATGACATTTATTTGTATCAAGACACCTTACTTCTGTGGGAGATTATCTGTGG 2139
Db 7204 TCATTTAGCAGATCGGCGAGAGACCACTTACTTCTGGGCTGAGATTCACTTTGGAAG 7263
Qy 2140 ACCGATTAACAGGACAGCCTGACAGATCGGAGACCTCTTAAAGGTTTTTAAGCTTG 2199
Db 7264 ATTGCTTACTCTCAGACGCGTGTGGGAGACCTTGAAGAGGTGTTTAAGTTGG 7333
Qy 2200 GAAAACATTGCGAGTGTATACCAAGACTCGACCGCGCGGCACTGATGATG 2259
Db 7324 GTAAACCGCTCCGACCGAGCATGAGCAAGACGAAGACAGAGACGCGCTCTGCTAATG 7383
Qy 2260 AAGCAATGCATGGAACAAGATTGGAATTACGAGATTAAGAGCGCTAGAAATCA 2319
Db 7384 AAACAAAGGCGTGGTTTAGATGATTAACAGACACTTATGAGTGGCGCTGGCAATC 7443
Qy 2320 GATACGATCATATCTGGACGCGCTGATCATCAGTCTCTGTCCACGTTAGCGAAGCG 2379
Db 7444 GGTATGAGTACACATCACTACCTGTCTGTGGCATGAGAACTTTTCCAGAGCA 7503
Qy 2380 TTAAAGACTTCAGAGCATTAAGAGAGCCCAATCACCTCTACGCGTGAACCTAAATAG 2439
Db 7504 AAAGAGACTTTCAAGCCATCAGAGGGGAAATTAAGCATCTCTACGCTGTCTTAATAGT 7563
Qy 2440 TGACGTATGAGA-----CAGCACTTACCCACCGGAGAAATTTTCA 2482
Db 7564 CAGCATGTATCTCATCTGACTAATATACCAACACACACCATGATGAGAGATTCT 7623
Qy 2483 TACCTCAGTGAACTTTCCACAGTTTACCTTCAATCCGATGGCTTACCGAGTCA 2542
Db 7624 TTAAATGCTTGGCGCGCGCTTCCAGCCCCACCTGCATGTGAGGCGCGAGAA 7683
Qy 2543 AACCTCTTACGCGCGCTGAGGCGCTTGGCGCGCGCTGCTCAATCGAAGAT 2602
Db 7684 GAGGAGGCGGCGCGCTGATGCTCCGCAATGGGCTGTCTCCCAATCCAGCACTGA 7743
Qy 2603 CTTAGAGGTGATGATCACTTCAAAAC--GATCACTTAATCGCGCCAG 2660
Db 7744 CCACAGCGGTACGTAGCTTGTGACAGGCAACTAGACTCAAAACCCACGCCAC 7803
Qy 2661 GTCCACGCGCAAGAGAGAGAGTGTCTTAGCCAAACCTTACAGCTTAAAGAA 2720
Db 7804 GCCCGCGCGCGCAGAGAGAGGCGCCAAAGCAACCGAAGCGGAAGAAACCA 7863
Qy 2721 AGAAGCAGCAAGCCAGAGAGCAAAAC---GCAAGCTTAAACAGAGAAACGAACGTA 2777
Db 7864 AAACACAGAGAAAGAAAGAGCACTGTCAAAACCCGGAAGAGAGAGCTGA 7923
Qy 2778 TGTGTATGAGTGGAGTGGACAAACATTTCCGATC--ATGCTGACGCGCAAGTGA 2834
Db 7924 TGGCACTTAAGTTGAGGCGCAGACTGTTCAGATCAAAAATGAGGACGGAATGCA 7983
Qy 2835 ATGATATGCTGTGGTGTGGAGAGGCTGATGAACCACTCAAGTTGAAGAAAA 2894
Db 7984 TCGGGACGCACTGGCCATGAGAGAAAGTAATGAACCACTCAAGTGAAGAACTA 8043
Qy 2895 TTGATTAAGCAATTAGCGCGCGTGAATTAAGAAAGGCTAGCATGTATGAGT 2954
Db 8044 TTGACACCGCTGTCTAACAAGCTCAATTCACCAAGTCTGACAGATACAGATGAGT 8103
Qy 2955 ACGGCGAGTTTCCCGAGAAATGAATTCAGACGCTGAGTACACGAGCAAAATCAC 3014
Db 8104 TCGCACAGTTGCGGCTCAACATGAGAGTGAAGGCTTCACTTACCTAAGCAACCTTG 8163

Qy 3015 CGGGCTTCTACAACTGGCACACGCGCAGTCCAGTATGAAATGGGAGATTTACCGTAC 3074
Db 8164 AAGGTTTCTACAACTGGCACACGAGCGGTGCAATATGAGAGAGATTTACCATCC 8223
Qy 3075 CGAGAGAGTGGCGGGGAAAGCGACACGGAAGACCGATCTTGACAAACAGAGCGAG 3134
Db 8224 CCGCGGATAGAGAGGACAGAGACAGTGTGTCGCAATTATGATTAATCTCAGGCGGG 8283
Qy 3135 TTGTGGTATTTGTTCTAGAGGTGCAATTAAGGGACAGCTAAGCGGCTTCACTGTGCA 3194
Db 8284 TTGTGGGATAGTCTCTGAGAGGGCTGATAGGGAACAAACCGCTTCTGCTGTCA 8343
Qy 3195 CTGGAACAGAAAGGGGATACCATTTAGGATACCCCGAAGTTTGAACCGGT--- 3251
Db 8344 CTGGAATACGAAGGGAAGACATCAAGCAACCCCGAAGGAGACGAAGAGTGTCTG 8403
Qy 3252 -----CACTAGTTACAGCGCTATGCGTCTTGAATGTCAAGTTCCATGGCAAAAC 3305
Db 8404 CTGACACACTGTGTCAGCGGCATGTGTCTTGGAACGTGAGCTTCCATGCAATGCC 8463
Qy 3306 CACCGTGTCTATTACTGACCGCAGAACGAACCTGACAGTCTGGAAGAAAGCTCG 3365
Db 8464 CGCCCATGCTACACCGGGAACATCCAGAGCTTCGACATCTTGAAAGAAACGTGA 8523
Qy 3366 ACAATCAAAATTACGACAGCTGTGGAAGACGTTGAAT---GTCCATACGCGCGC 3422
Db 8524 ACCACAGGCTTACGACACCTGTCTCAAGCCATATTTGGGTGCGATCTCTCGGCAAA 8583
Qy 3423 CAAAAGCAAGCATTTACCGATGACTTCACTGACCAAGTCTTACCTTGGGTTTGCCTG 3482
Db 8584 GTAAAGAGACGTCACTGACGACTTATCTTGAACAGCCGTAATTGGGCAATGCTGT 8643
Qy 3483 ATTGCAACACTCAAGCGGTGTTCAGCCCAATAAATTTGAGAAAGTGTGGAGCAAT 3542
Db 8644 ACTGTCAACATCTGAACCGTGTCTTGAACCGATTAAGATCGAGAGCTGTGGATGAAG 8703
Qy 3543 CTGATGATGATGATGATGAAATCCAGGCTTCGGCACAATTCGGTCAATCAAGCAGCA 3602
Db 8704 CGGAGCAACACCATACGATACAGACTTCCGCGCATTTGGATAGACAAAGCGGAG 8763
Qy 3603 CTGCGATGTCAACCAAAATTCGTTACATGTCTTTGACACAGACCATGATCAAGGAG 3662
Db 8764 CAGCAAGCTCAAAATTAAGTACCGGTACATGTCTGAGAGAGATCATATGTCAAAAG 8823
Qy 3663 ACAGTATGAGAAATATGCTATCAGCACTGTGACCTTCCGTCTTGGCCAAAG- 3722
Db 8824 GCACATGATGATGATCAATCAAGATCAGACCTCAGAGACGTTAGAAAGCTTACAAAG 8883
Qy 3723 GGTACTTCTGTAGTCAATGTCTCTCAGGTGACAGTGAACGTCAGTATCAGAGCG 3782
Db 8884 GATATCTTCTCTCGGCAAGTGTCTCAGGGAGACGCTTACGATAGCGAGTA 8943
Qy 3783 GAGCATGTAGAAATTCATGACCGGTGAGAAAAAGATCAGAGAGAAATTTGTGCTAG 3842
Db 8944 GCAATCAGCAAGTATGACATGCAATGCGCGCAAGATTAACCAAAATTCGTGGAGCGG 9003
Qy 3843 AGGAGTACTTGTTCACACCGCTCATGAAAGCTGTGAAGTGCACAGTTTACATCACT 3902
Db 9004 AAAAATATGACCTTACCTCCGTTCAACGTTAAGAAATTTCTTGAACAGTATGACACGTC 9063
Qy 3903 TGAAGAGAGCTGCGCGGGTACATACATGACAGAGCCAGGCGCACAGCGTATAGT 3962
Db 9064 TGAAGAGAAACACCGCGGCTACTATCTATGCAAGGCGGGAACGCAAGCTTATCAT 9123
Qy 3963 CTTATCTGAGAGAGCTGACGAGGAACTGTATTAACCACTTCTGGAGAAAGCTCA 4022
Db 9124 CTTATCTGAGAGATCATTCAGGAAAGTTTACGGAAGCCACCATTCGGGAAAGAACATTA 9183
Qy 4023 CTTACGAATTAAGTGTGGGACTACAGCACAGTATCTGAGACCGGAAACGAAGTGA 4082
Db 9184 CTTACGAGTGAAGTGGGAGATTACAAAGCCGGAACCGTTATGACACCGAATCA 9243

Db 11404 GCATATGATCAGCTGCTATATTAGATCCCGCTTACCGGGGCAATATAGCAAGACC 11463
Qy 6269 AAACTGATGATTTTCCGAGAGCAAGTGCATTAATGCTGTCAGT 6318
Db 11464 AAACTGAGCTATTTCCGAGAGCAAGTGCATTAATGCTGTCAGT 11513

RESULT 8
AR042409
LOCUS AR042409 11663 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5811407.
ACCESSION AR042409
VERSION AR042409.1 GI:5962905
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 11663)
AUTHORS Johnston,R.E., Davis,N.L. and Simpson,D.A.
TITLE System for the in vivo delivery and expression of heterologous
JOURNAL Patent: US 5811407-A 1 22-SEP-1998;
FEATURES Location/Qualifiers
source 1..11663
/organism="unknown"
BASE COUNT 3291 a 3105 c 2907 g 2360 t
ORIGIN

Query Match 36.9%; Score 2396; DB 6; Length 11663;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3751; Conservative 0; Mismatches 1985; Indels 74; Gaps 9;

Qy 580 ACTGACGGTATGAGCGGAGCGGTATATTTCTCATGGAAGACGGCCAAAGTCACTTC 639
Db 5707 ACTGTCACCGGGATAGTGGTACATTTTGCAGCAGACAGCCCTGGCACTTC 5766
Qy 640 AACGAAATCAGTACGTCAATGTAACACAGAACCTATATGATCGGCGCTCATG 699
Db 5767 AAAAGAGTCGCTTCTGAGAACAGCTTACAGAACGACCTTGAAGCGAATGTTCTG 5826
Qy 700 AGAAGTATACGCGCGCGCTCATCTCGAAGAGAAATGTTACAGAAACTGC 759
Db 5827 AAACATATCAAGCCCGCTGCTGCACAGTCGAAAGAGAACAGCTCAACTCAGTACC 5886
Qy 760 AATATGCGCTCTGAAGAAATAGAGAGATATCATACGAAAGTAAATATGA 819
Db 5887 AGATGATGCCACCGAAGCAACAAAGCAGTACAGTCTGAAAAGTAAACCAAGA 5946
Qy 820 AAGCAATTAACAGCGAGCACTATTTCTGATTTGGGACATATCTATCATCAGAAGTA 879
Db 5947 AAGCATTAACCACTGAGGAGCTGTTTCAAGGCTTACAGCTGTATTAAGCTTGCACAG-- 6063
Qy 880 ATCTGTGAGTGTATACAGATCAATTAATCTGTACCAATTAATCTGTCACAGTAATTA 939
Db 6004 ATTCAGCCAGATGCTATAGATCACTAACCCGAACCACTGATTTCCAGCAGTATCCAG 6063
Qy 940 ACAGGTTTAACTGCGAGAGTGGCGGTTAAAGTGAATTAAGTTATCCAGAGAT 999
Db 6064 GAACTACTCTGACCCAAAGTTTGTGAGTGTGTTGTAACACTATCTGATATGAATTT 6123
Qy 1000 ACCCTAAGTACGACGATATATGTAATACAGATGAATAGATGCTATCTGACATGTTG 1059
Db 6124 ACCGAGCGTATGATCTTATAGATCAACGAGATGATGCTTACTTGATATGTTAG 6183
Qy 1060 ACGGCGATCTGCTGTCTAGATACAGCACTTTTGTGCGGCTAACTGAAAGTACC 1119
Db 6184 ACGGAGAGTGTGCTTATGATATCTGCACTTTTGTGCGCAAGCTTGAAGTTACC 6243
Qy 1120 CAAGAAGCATATGTTTGAAGCAAGATATAGATACGCGTCCATCGCTATACGA 1179
Db 6244 CGAAAAGACAGATATAGAGCCCAACATCCGAGTGGGTTTCATCAAGCATGACGA 6303
Qy 1180 ATACATTACAAATGTAATTTGGCTCAGCTACTAAAGGAATTGCAACGTTACCAATTC 1239

Db 6304 ACAGTTGCAAAAGCTCATTTGCGCGACTAAAGAAATGCAAGCTACACAAATGC 6363
Qy 1240 GAGATTAATCTGCTTATGATTTCCGCGGCAATTTATGTTGATTTTCAAGAAATACCAT 1299
Db 6364 GTGAATCGCAACACTGAGACTCAGCGCAATTAACGTTGAATGCTTTCGAAAATATCAT 6423
Qy 1300 GCAATGATGATGACTGGGATACCTTTGCGATTAACCTATTTGGCTTAATACAGAGACG 1359
Db 6424 GCATATGAGATATTTGGAGAGAGATTTCGCCGAAAGCCAAATTAGATACATAGATTCG 6483
Qy 1360 TTACGCAATATGTCACAAAGCTGAAGAGCGGCAAGAGAGACGATTTTTCGAAATCTC 1419
Db 6484 TTACCGCATAGTGGCCAGACTGAAGAGCCCTTAAGCGCGCCGCACTGTCGAAAGACGC 6543
Qy 1420 ATATATCAAAACGCTTGCAGAGATATCAATGATCAATTCCTCATGATCTTAAAGAG 1479
Db 6544 ATATATTTGTCCATTCGCAAGAGTCCCTATGATATTCCTATGACATGAAAGAG 6603
Qy 1480 ATGTCAAAGTTATCCCGGACAGCAATACATACAGAGAGCGGCTTAAGTCAAGTATTC 1539
Db 6604 ACGTGAAGTTACACTCTGGACAGAAACACAGAAAGAAACCGAAAGTACAGTATAC 6663
Qy 1540 AAGCTGCAATCCCTTGTCTACCGCTTACCTTTGCGGAGTTCATCGGAAATTAAGTCGTA 1599
Db 6664 AAGCCGCAAGACCCCTGCGACGCTTACCTATGCGGAGTTCACCGGAGTATAGTCGCA 6723
Qy 1600 GACTGAATGGGCTCTTCTGCAAAATTCATATCTCTTTCGACATCTCAGCGGAAAT 1659
Db 6724 GCGTTAAGCGGCTTGTCTACCAAACTTACACAGCTCTTTGCAATATCGCGGAGAGACT 6783
Qy 1660 TTGATGCAATATTTGTCGAACATTTCCACAGCGGACCAAGTATGGAAGCGACATCG 1719
Db 6784 TTGATGCAATATTTGTCGAACATTTCCACAGCGGACCAAGTATGGAAGCGATATCG 6843
Qy 1720 CGTGTGTAATTAAGGGAAGACGACGCTATTCGCAATTTGCGGCTGATGATCTTTAGG 1779
Db 6844 CTTGTTGCAAAAGCGCAAGACGACGCTATGCGGCTTAAACCGGCTGATATCTTTGAGG 6903
Qy 1780 ACTTATGCTGACCAACCGCTCTTATGATTTGATGAGGGGGGCTTGGCAATTCAT 1839
Db 6904 ACTGGGTGTGACCAACCACTACTGATTTGATGAGTGGCTTTTGGAGAAATATCAT 6963
Qy 1840 CTGTGACCTACCTACAGAAAGAGTTTAAATTTGGTGCATGATGAAATCCGATATG 1899
Db 6964 CCACCATCTGCGCAGGGATACCGCTTCAATTTGGGGGATGATGAATCCGATATG 7023
Qy 1900 TCTTAACGCTGTTTGTCAACACACTAGTCAATATCATGATTTGCTAGCAGATCTACG 1959
Db 7024 TCTCAGCGCTCTTTGTCAACACAGTTCTGAATGTCGTTATCGCAGAGATTTGAGG 7083
Qy 1960 AAGCGTTAACACGCTGACGCTGCGCGGCTCTATTCGGCGAGATTAATATAGTGCATG 2019
Db 7084 AGCGGCTTAAACGCTCAAAATGTGACACATTTATGCGAGCAACATTAATTAACGAG 7143
Qy 2020 TCGTCTCCGACACCTTATGCGGAGAGATGCGCACTTGGCTGAACATGAAAGTAAATA 2079
Db 7144 TAGTATCTGACAAAGAAATGCTGAGAGTGTGCACTGGCTCAACATGGAAGTTAAGA 7263
Qy 2140 ACCAGATTAACAGGACAGCTGCAAGATGCGCAACCTCTTAAAGAGCTTTTAAGCTTG 2199
Db 7264 ATTGCTTACCTCCACAGCGGTGTGCTGCGGACCTCTTGAAGAGCTGTTTAAGTTGG 7323
Qy 2200 GAAATACATTTGCGGATGATATCCCAAGACTGCGACCGCGCGGCACTGCATGATG 2259
Db 7324 GTAAACCGCTTCCAGCGGACGATGAGCAAGACGAAGACGCGCTCTGCTATATG 7383
Qy 2260 AAGCAATGCGATGGAACAGATTTGAATTAACGACGATTAAGTAAAGCCGTAGATATCA 2319

Dh 7384 AACAAAGCGCTGTTAGAGTAGGTATPACAGACACTTGTAGCAGTGGCCGTGGCACTC 7443
Qy 2320 GATACGATCATATACTGGCAGGCTGTATCATCCTCTCTGTCAAGTTAGCCGAAGG 2379
Db 7444 GGTATGAGTAGACAAATCACACCTGTCTCTGTGGAATTTTGGCCAGAGCA 7503
Qy 2380 TTAGAACCTTCAAGAGCATPAGAGGAGCCCAATCACTCTTACGGCTTGAACCTTAATAG 2439
Db 7504 AAAGAGCATTTCAAGCCATCAGAGGGGAAATTAACATCTTCAAGGTGCTTAATAGT 7563
Qy 2440 TGAGCTAGTAGA-----CAAGCACTTAACCAACCGGAGAAATGTTTCCA 2482
Db 7564 CAGCATAGTAGCTTTTCATCTGATTAATCACACACACACATGATTAAGATTTCT 7623
Qy 2483 TACCTTCAGCTGAATTTTCCACCACTTAACTCTAATAATCCGATGGCTTACCGAGATCCA 2542
Db 7624 TTAACATCTGGGCGGCGGCGGCTTCCAGCGCCCACTGCAATGTGAGAGCGCGGAGAA 7683
Qy 2543 AACCTCTTAGGCGCGCTGAGAGCGTTTCGGCGCCGCTGGCTGCTCAATCGAAGAT 2602
Db 7684 GGAGGCTAGCGGCGCCGATGCTGCGCCGCAATGGGCTGCTTCCCAATTCAGCAACTGA 7743
Qy 2603 CTTAGAGAGTGCATAGTCACTTGACTTTCAAAACAC--GATCACTTAATCGCGCCAG 2660
Db 7744 CCACAGCGGTGACGCTTACCTAGTCACTTGAACAGCACTAGAACCTTCAACCCCAACG 7803
Qy 2661 GTCCACCGCCCAAGAAAGAAAGAGTGTCTTAACTCAAACTCTACGCTTAAAGAA 2720
Db 7804 GCCCGCGCGCGCCGCAAGAAAGAGAGCGCCCAAGCAACCCGAAAGCCGAAGAAACCA 7863
Qy 2721 AGAAGCAGCAAGCCAAAGAGAGCAAAAC--GCAGCTTAAACAGAGAAACGACAACTGA 2777
Db 7864 AAACACAGAGAAAGAAAGAAAGCACTGCAAAACCCAAACCCGAAAGAGACGCGTA 7923
Qy 2778 TGTGTATGAATTTGAGTTCGACAGACATTTCCGATC---ATGCTGAACCGCCAGTGA 2834
Db 7924 TGGCATTAAGTTGAGGCGCGACACATGTTTCAGCTCAAAATGAGAGCGAGATGTCA 7983
Qy 2835 ATGATATGCTGCTGTGCGAGAGAGCTGATGAACCACTCCACGTTGAGGAA 2894
Db 7984 TCGGCGACGCACTGCGCATGGAAGAAAGTATTAACCACTCCACGTTGAAGGAACTTA 8043
Qy 2895 TTGATATGAGCAATTAGCGGCGGTGAATTTGAAGAGCTAGCATGTACGACTTGGAGT 2954
Db 8044 TTGACACCTGTGCTATCAAGCTCAATTCACCAAGTCCGACATACGACATGAGT 8103
Qy 2955 ACGGCGAGCTTCCCAAGAACTGAATACAGACCTGCAATGACACCGCAAAACCA 3014
Db 8104 TCGACACATGTGCGGTCAACATGAGAGTGAAGCGCTTCACTACACAGTGAACACCTG 8163
Qy 3015 CGGCTTCTACCACTGGACACCGCGGAGTCCAGTATGAGAAATGGAGATTTTACCTTAC 3074
Db 8164 AAGGCTTTACCACTGGACACCGAGCGGTGCAATATGAGAGGCAAGATTTACATTC 8223
Qy 3075 CGAGAGAGTGGCGGGAAGAGCGCAACGCGAAGACCGATCTGTGACAAACAGAGCAG 3134
Db 8224 CCCGCGAGTAGAGGAGAGAGACATGTGCTCCATTAATGATTAACCTCAGCGCGG 8283
Qy 3135 TTGTGCTATTTGTTTGAAGAGTGAATGAGGGCAGCGGTACGGCGCTTTCAGTGTCA 3194
Db 8284 TTGTGCTCATGTCTCTCGAAGGGGCTGATGAGGGAACAAAGCCCTTTTCGCTGCTCA 8343
Qy 3195 CTGGAACCAAGAGGGGTGACCATTAGAGATACCCCGAAGGTTCTGAACCGTGT--- 3251
Db 8344 CTTGAATATGCAAGAGGAAAGAAATCAAGACACCCCGAAGGACAAAGAGTGTCTG 8403
Qy 3252 -----CACTAGTTACAGCGCTATGCTGCTTTCAATGTCACTTTCCATGCGCAAC 3305
Db 8404 CTGACACACTGTGTACCGGCACTGTGCTTGTGAAAGTAGAGCTTCCATGCAATGCC 8463
Qy 3306 CACCGGTGTCTTATCTAGACCGCAAGACCACTGAGGTGCTGAGAGAAAGTGTG 3365
Db 8464 CGCCCACTGTCTACACCGCGAAACCATCAGAGCTCTGACATCTCTGAAGAGAACTGA 8523

Qy 3366 ACAATCCAAATTACGACACGCTGCTGAGAAAGTCTTGAAT---GTCCATACCGCCGC 3422
Db 8524 AOCAGAGCGCTACGACACCTGTCTCAACGCCATATTCGGGTGCGGATGTCTCCGCAAA 8583
Qy 3423 CCAAACGAGCATTTACGATGACATTTACATGACCACTTCCCTTACCTGGGTTTCCCGT 3482
Db 8584 GTAAAGAGGCGTCACTGACATTTTACCTTGAACGAGCCGTAATTTGGGACATGCTCGT 8643
Qy 3483 ATTGACGACTCAACGCGGTGTTCAAGCCCAATTAATTAAGAAACGTGTGGGCGAAT 3542
Db 8644 ACTGCAACATATGAAACGCTGCTTACCTCGATTAATGATGAGGAGGTGTGGATGAAG 8703
Qy 3543 CTGATGATGATTCATTTGAATTCAGAGTCTGGGCAATTTGGCTTCAATCAGGAGCA 3602
Db 8704 CGAGAGCAACACCATACGATACGACTTCCGCCAATTTGATACGACCAAGCGGAG 8763
Qy 3603 CTGGGATGTCAACCAATTCGTTACATGTCTTTGACACGACCAATGACATCAAGAG 3662
Db 8764 CAGCAAGCTCAAAATTAAGTACCGTTACATGTGCTGACAGGATCATCTGTCAAGAG 8823
Qy 3663 ACAATATGAGAAATATGACTATCAGCACATTTGACCTGCGGTGCTTGGCCCAAG 3722
Db 8824 GCACATGATGACATCAAGATCAGCACCTGAGACCGTGTAGAAAGCTTAGCTTCAAG 8883
Qy 3723 GGTACTTCTGTTAGCTCAATGTCTCTCCAGGTGACAGTGTAAACGCTTATCAGAGCG 3782
Db 8884 GATACTTCTCTCGGCAAGTGTCTCTCCAGGGACAGGTAACTGATAGCAATG 8943
Qy 3783 GAGCATCTGAATTCATGACACGCTGAGAAAGAAAGATCAGAGAAAGTTTGTGGTAG 3842
Db 8944 GCAACTCAGCAACGTATGACCAATGAGCCGCAAGATTAACAAATTTCTGGGACGG 9003
Qy 3843 AGAGTACTTGTTCACACCGCTCCATGGAAGCTGTAAATGCTCACTTTACATCACT 3902
Db 9004 AAAATATGACTTCTCTCCGTTACAGGTAAAGATTTCTTGAACATGTACGACCTG 9063
Qy 3903 TGAAGAGACGCTGCGGGTACATTAACCATGACACAGGCCACAGCGTAAAGT 3962
Db 9064 TGAAAGAAACACCGCGGCTACATCACTATGACAGGCCGAGACCGATGCTTATCAT 9123
Qy 3963 CCTATCTGGAGAAACGCTCAGGCGAAGTATCATTTAAACCACTTTCGCAAGAACGTC 4022
Db 9124 CCTATCTGAGAGATTCATCAGGAAAGTTTACGGAAAGCCACATCCGGAAAGAACATTA 9183
Qy 4023 CTTACGATTTAATGTGTGCGCACTACACAGAGTATGTGTAGACACGGGAACGAATGA 4082
Db 9184 CGTACGAGTGCAGTGTGCGGCAATTACAAAGACCGAACCCTTACGACCTGACCGAAATCA 9243
Qy 4083 ACGGTGCACTAAAGCAAAACAGTGCATTTGCTCAAGAGGAGCAAAAGGAGTCT 4142
Db 9244 CGGCTGACCGCCATCAAGAGATGCTGCTTTAAGAGGACCAAAAGAGTGGTCT 9303
Qy 4143 TCAACTGCGCGGATCTTATTAAGCAACAGACCACTCAGTGCAGATTAATTTGACATTC 4202
Db 9304 TCAACTGCGCGGATCTGATGACAGACCGACCAACGCGCCAAAGGAAATTTGACTTGC 9363
Qy 4203 CATTCGCTTACACCGACATGTGCGCGGTTCCGTTAGCTACACGCTTACAGTACGA 4262
Db 9364 CTTTCAAGCTGATCCGAGTACTGTGATGTGCTGCTTGTGCGCACGCGCGAAGTATAC 9423
Qy 4263 AGTGGTTCAAGGATCAACCTTCACGATGACGAAATGCAAAATTTGTGCAACAG 4322
Db 9424 ACGCTTTAAACATATGCTTCCAAATTGACACAGACCATGTGACATTTGCTACACCA 9483
Qy 4323 GAAATTTGGGCTCGAGACAGCAACAGCAGATGATTAAGGCTTACATCTCAGGA 4382
Db 9484 GGAAGCTAGGGGCAAAACCGGAAACCACTGATGATGATCAATGGAACCGGTTGAA 9543
Qy 4383 ATTTTCTGTGGGCGAGAGAGGCTGAGATTCGATGGGATTAACATGAACCACTCAG 4442
Db 9544 ACTTCAACCTGACCGAGATGCGCTGGAATTCATATGGGCGCAATCAAGAACAGTAAAGG 9603

QY	4443	TCCTGGGCGCCAGAGTCCGGCACACAGCGCCACCACTGGATGGCCGGATAGATCAATCAATCC	4502
Db	9604	TCCTATGCCCAAGAGTCTGCACACAGAGACCCCTCAGAGATGGCCACAGAAATAGTACAGC	9663
QY	4503	ACTATATATCATCGGCAATCCAGTCTTCACTGTCATATGTGCTGTGTGTGTGCTCTGTGCTA	4562
Db	9664	ATTACTATCATTCGCGCATCTGTGTATACCATCTTAGCCGTGCATCAAGCTCTGTGGCGA	9723
QY	4563	TCTCTGGTAAAGGCACTGCATCATACGAGCTTGCAATGCGAAGCAAGAAAGACCTGCCTGA	4622
Db	9724	TGAATGATTTGGGTAACTGTGTGACAGCATTAATGTGCTGTAAAGCGCGCGTGAAGTGCCTGA	9783
QY	4623	CGCCATACGCGCTTGCACCGAAGCAACGATACCGACACAGCATTTGAACGGTTTGTGTGCTCA	4682
Db	9784	CGCCATATGCCCTTGCCGCCCAATATGCCGATTTCCAACTTCCGTGCGACTTTTGTGTGTG	9843
QY	4683	TTCCGGCCAAACCAAGCTGTAAACATTTTGGAGAACTTTGAAACCATCTGTGTATTAACAAC	4742
Db	9844	TTTAGTGGGCTAAATGTCTGAACATTCACCGAGACATATAGTATCTTATATGTGTGAACGCC	9903
QY	4743	AAACGTTTCTGTGGGCACAGTTGTGCATTTCTGTGGACGCCCTTGTTATTTCTGTCCGCT	4802
Db	9904	AGCGGTTCTTGTGGGTTCAGCTGTGTATACCTCTGGCCGCTGTGCTGTCTTAATGTGCCT	9963
QY	4803	GCTTTTAACTGTGCATGCGCTTTTATTTATGTGTGCAGCGGTCTGCTGGGGAAAGTAAAGC	4862
Db	9964	GTTCTCTCAATGTCTGCTGCTCTTTTATTTAGTGTGTCGCGGCGCTTACTGTGCGAAGGTAAAGC	1002
QY	4863	CCTTTCGAACATGCAGCACCACTGTGCCTGAATGTTCGGGGATCCCGTATTAAGACGTTTGTGC	4922
Db	10024	CCTTAGAACAATGCAGCACCACTTTCCMAATGTGTGCCACAGATATCCGATTAAGGCACTTGTGTG	1008
QY	4923	AAACGCGCAGTTACGCGCCACTTAACCTGTAGATCACGCTGTCTCATCGGAATTAACAAC	4982
Db	10084	AAAGGCGCAGGTTACGCCCCGCTCAATTTGGAGATTACTGTATGTCTCGAGGTTTGTGC	1014
QY	4983	CTTCAACTTAACAAGAGTACGTGACCTGCAATTCACACAGTATTCCTTACACACAG	5042
Db	10144	CTTCCACCAACCAAGAGTACATTAACCTGTCAAAATTAACCATCTGTGTCTCCCTCTTAAG	1020
QY	5043	TTTAATGCTGCGGGTCCCTCGAGTGCAGAAGGATCCTCAAAAGCGGATTAATACATGCGCG	5102
Db	10204	TCAATGTGTGGGCTCTCTTGGAAATGTCAAGCCGCGCTCAGCGACATTAACCTGTCAAG	1026
QY	5103	TTTTTGGCGGTGTATCCCTTTTCATGTGTGGGAGGCGCAATGTCTCTGTGACAGTGA	5162
Db	10264	TCTTTGGAGGGGTGTACCCCTTTCATGTGTGGAGGAGCAATGTTTTGGCAGACGTGA	1032
QY	5163	ACACACACACTAGTGAAGGCGTACGTTCGATTCGCTCCAGACTGCACATTAATACAGCAG	5222
Db	10324	ACACCCAGATTAAGTGAAGGCGTACGTTCGAAATGTCAATGATTTGGCGCATACACACGCGC	1038
QY	5223	TTCGCACTAAAGTTTCAACACAGCTGCTCGAAGTTCGAGGCGCTCGATATAGTATACGGCAAC	5282
Db	10384	AGCGGATTTAAGGTGTATCTGCCGCGATGAAGTGAAGACTCGCTGATATGTATGACGGGAACA	1044
QY	5283	CCACCGCGCACCTGGATACGTTTGTTCATATGGCGTCAACGCCAGTGTCTCAACGGGACCTGA	5342
Db	10444	CTACACAGTTTCTTATGTGTATGTATGTAACGAGATACACACAGGAACGCTTAAAGACTGA	1050
QY	5343	AGGTCTATAGCAGGGCCGATATACGCCGCTTTTTCACCTTTGACCATTAAGTGTGTATCA	5402
Db	10504	AAATCATATAGCTGCACCAATTTACAGATTTGTATTAACCTTGTGATTCACAAAGTGTGTATCA	1056
QY	5403	GAAAGGGGCTGTTTATCAACTACACTCCCTGAGTATGAGGCTATGAACACAGAGAGCT	5462
Db	10564	ATCCGGGCTGTGTACAACTATGACTTTCCGGAATATGAGGACGATGAACACAGAGAGCT	1062
QY	5463	TCCGCGATATTTCAAGACATCTCGCTGTATGTCTACAGACATATAGTATGCGCCGACTGACATAC	5522
Db	10624	TTGGAGACATTTCAAGCTACTCTCTTGTATAGCAAAAGACTCATATGCCACGACACAGACATTA	1068
QY	5523	GGCTGTGAAGCTTTCTGTCAAGAACATCACTGCCCTTACACCCCAAGACAGTATACGGT	5582

Db	10684	GGCTACTCAAGCCTTCGGCCAAAGAACTGCATGTCCGTTACAGCAGAGCCGCATCTGGAT	1074		
Oy	5583	ATGAATGTGAGAACAACTCAGAGACGACCCCTGCAGAAACAGACCAATTTGGATGTA	5642		
Db	10744	TCGAGATGTGGAAAAAACAACTCAGGCCGCCACCTGCAGGAAACCGCCCTTTTGGGGTCA	1080		
Oy	5643	AAATTGAAGTGGAGCCCTCTGCGAGCCGTCTAACTGTGTCTTACGGGACATACCCGATCTCGA	5702		
Db	10804	AGATTGGACGTCAATCCGCTTCGAGCCGGTGGACGTGCTCATACGGGAACTTCCCATTTCTA	1086		
Oy	5703	TTGCATCCCTGATGACGCTTTTGTGAGATCATCAGAAATCACCMAAATTTAGAAATTA	5762		
Db	10864	TTGACATCCGGAAGCGCTGCTTTATACAGACATCAGATGACACCACTGCTCTCAACAGTCA	1092		
Oy	5763	GCTGCACAGTACGACAGCTGCATTTATTTCTGCAAGCTTTGGTGTCTCTTAACTTACAGT	5822		
Db	10924	AATGTGATGTACAGTGAATGACCTTATTCACGCGGACTTCGAGGGATGGCTTACCTCGAGT	1098		
Oy	5823	ACAAAGCTGACAGGAGGAGACATTTGTCACTTCACTCCCACTCCACAGACAGCTGTTTGA	5882		
Db	10984	ATGTATCCGACCCGGGAAGCAAAATGCCCTGTACATTCGACTTGCAGACAGCAACCTCC	1104		
Oy	5883	AGAAAGGACACACATGTGACTGCGGTAGGACAGACATPAACATACTTTTAGACATCGA	5942		
Db	11044	AAGAGTGACAGCTTCATGTCTCTGGAAGAAAGAGCGGTGACAGTACACTTCAGCACCGCGA	1110		
Oy	5943	GCCCAACAGCAAAATTTTATATAGTTTCGCTATGCGGCAAGAAAGTCCACCTGCATGTCTGAAT	6002		
Db	11104	GCCCAACAGCGAACTTCATGTATCGCTGTGTGGTAAAGAACAAATGCATGCAATGCAGAAAT	1116		
Oy	6003	GTAAACCAACGCCCGACCACTAATTTGGAAACACATTAAGTCAGACCAAGAAATTCAGG	6062		
Db	11164	GCAAAACCAACAGCTGATCATATCGTGGACACCCCGCAAAAAATGACCAAGAAATTCAGG	1122		
Oy	6063	CGGCAAGTTTCCAAAACATCTTGGAACTGCGCTGCTTGCACTGTTTGGGGAGACATCATCC	6122		
Db	11224	CGGCATCTCAAAAACATTCATGAGAGTTGGCTGTTTGGCTTTTCGGGGGCGCTCGTCGC	1128		
Oy	6123	TCATTGTTGAGACTTATATAGTTGTGCTGACAGCTTATGCTTATTAACACAGTAGAT	6182		
Db	11284	TATTAATTATAGACTTATATGATTTTTCCTTGACAGATAGTGTGACTAGACACAGAAAGAT	1134		
Oy	6183	GACTGAGCGCGGACACTGA-CATPAGCGGTAAACCTCATGTACTTCCGAGAGAACGTGT	6241		
Db	11344	GACCGCTACGCCCCAATAGACCCGACCCGACAAAACTCGATGTACTTCCGAGAACTGATGT	1140		
Oy	6242	GCATPAATGCCAGCGCGCGCT-----TGACACT	6268		
Db	11404	GCATPAATGCATCAGGCTGTGTATATTAGATCCCGCTTACCGCGGCAATATAGCAACACC	1146		
Oy	6269	AAAATCGATGATTTTCCGAGAGAGCAACAGTGCATATATGCTGTGACAGTGT	6318		
Db	11464	AAAATCGAGATTTTCCGAGAGAGCGCAGTGCATATATGCTGTGCGAGTGT	11513		
RESULT 9	AR096555	11663 bp	DNA	linear	PAT 08-SEP-2000
LOCUS	AR096555	Sequence 1 from patent US 6008035.			
DEFINITION	AR096555				
ACCESSION	AR096555				
VERSION	AR096555.1	GI:10025463			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 11663)				
TITLE	Johnston, R.E., Davis, N.L. and Simpson, D.A.				
FEATURES	System for the in vivo delivery and expression of heterologous genes in the bone marrow				
REMARKS	Patent: US 6008035-A 1 28-DEC-1999;				
LOCUS	Location/Qualifiers				
LOCUS	1. 11663				

BASE COUNT 3291 a 3105 c 2907 g 2360 t
ORIGIN
Query Match 36.9%; Score 2396; DB 6; Length 11663;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3751; Conservative 0; Mismatches 1985; Indels 74; Gaps 9;
/organism="unknown"
580 ACTGACGATGAAACGCGGAGCTATTTCTCATCGGAAACAGGCCAAGCTCACTTC 639
Db ACTGCTAACCGGGGTAGGTGGTACATATTTTCACGACACAGCCCTGGGCACTTGC 5766
Qy 640 AACAGAAATCAATGCTCAATGTAACTACAGAACTATATTTGATGCGGCGCTCAATG 699
Db 5767 AAAAGAAATCGGTTTGCGAAGACGAGCTTACAGAACGACCTTGAAGGCGAATGTTCTGG 5826
Qy 700 AGAAGTATTTAGCCCGCGCTCGATCTCGAAAGAGAAATGTTTACAGAAAGAACTGC 759
Db 5827 AAAGATCTACGCGCCCGGTGCTCGACAGTCCGAAAGAGAGAACGCTCAAACTCAGGTACC 5886
Qy 760 AATTATGCGCCTCTGAGAAATAGAGAGGTATCATCTCCGAAAGTAAATATAGA 819
Db 5887 AGATGATCCCAACGAGCCCAACAAAGAGGTACAGTCTCGAAAGTAAAGCAAGCA 5946
Qy 820 AAGCAATTACAGCGAGGAGCTCATTTCTGATTTGGGACATATCTATCATCAGAACTGA 879
Db 5947 AAGCATTAACCACTAGGAGCTGCTTTCAGGGCTACGACTGATTAACCTGCGACAG--- 6003
Qy 880 ATCTGCTGAGTGTTCACAGAGTCAATTTCTGTACCAATCTACTCTGCAACGCTAATTA 939
Db 6004 ATCAGCCGAAATGCTATTAAGATCACCTACCCGAAACCACTGATTTCCAGAGTATCACAG 6063
Qy 940 ACAGGTTTACATCTGCAAGAGTGGGTTTAAACGTGCACTTATTTTCAAGAAATT 999
Db 6064 CGAATCTCTGACCCCAAGTTTGTGTAGCTGTTTGAACAATATCTGATGAGAAATT 6123
Qy 1000 ACCCTACAGTACCGAGTATTGTATACAGATGAATACGATCGTATCTTGACATGTGG 1059
Db 6124 ACCGACGCTGACATCTTATTCAGATCACCGAGATGAGATGCTTATTTGATATGTGAG 6183
Qy 1060 ACGGCGCATCTGCTGTCTAGATACAGCCACTTTTGTCCGGCTAAACTGAGAACCTTACC 1119
Db 6184 ACGGACAGTCTGCTGTCTAGATACAGCCACTTTTGTCCGGCTAAACTGAGAAATTACC 6243
Qy 1120 CAAGAATACATAGCTATTTTTCAGCCAGATGAATACAGCGTCCCATCTGCTATACGA 1179
Db 6244 CGAAAGACACGAGATATGAGAGCCCAACATCCGAGTGGCTTCCATCAGAGATGACGA 6303
Qy 1180 ATACATTCACAAATGTATTTGGCTGAGCTACTTAAAGGAATTCGAACGTTACCCAAATGC 1239
Db 6304 ACACCTTGCMAAAGCTGCTCATTTGCCGAGCTTAAAGAAATCTGCAACGTCACACAAATGC 6363
Qy 1240 GAGAAATTACTGTCTTGAATTTGGCGGCAATTAATGTTGATTTTCAAGAAATACGAT 1299
Db 6364 GTGAATCTCCAAACATGACTCAGGACATTCAGCTTGAATGATCTTTGAAAATATGAT 6423
Qy 1300 GCAATGATGATGATCTGAGATACCTTTCCGAGATACCCATTTGCGCTAACTACAGAAAGC 1359
Db 6424 GCAATGAGAGATTTGGAGAGATTTGCCGAAACCAATTAAGATCTCTATGTAGTTTCG 6483
Qy 1360 TTACGCAATATGTGCAAAAGCTGAAGAGGCGGAAAGCAGACAGCTTGTTTGGGAATATCTC 1419
Db 6484 TTACGCGATACGTGGCCAGATCTGAAGAGCCCTTAAGGCCGCGCATGTTCCGAAAGAGCC 6543
Qy 1420 ATTAATCTAAAACCGTTGACAGAGATACCAATGATCAATTTGCTCATGATCTTAAAGAGAG 1479
Db 6544 ATTAATTTGCTCCATTTGCAAGAAAGTCCATATGATAGATTTGCTCATGACATGAAGAGAG 6603
Qy 1480 ATGTAAAGATTAATCTCCGCGACGAAACATACAGAGAGCGGCTTAAAGTGCAGGTTATTC 1539
Db 6604 ACGTAAAGATTAACACTCTGCGACGAAACACACAGAAAGAAAGCCGAAAGTATCAAGTGAATAC 6663
Qy 1540 AGGCTGACAGATCCCTTGCTACCGCTTACCTTTGCGGGATTCATCGGGGAATTAATGCTCGTA 1599

Db 6664 AAGCCGACGAACCCCTGGCGACCGGCTTACTATGCGGATCCACCGGGAGTTAGTCCGA 6723
Qy 1600 GACTGATGCGGTGCTTTGCGCAATATCCATCTCTTGGACATGTCAGCGGAAGATT 1659
Db 6724 GGCCTTACGCGGCTTTTGTACCAACATTCACACGCTTTTGAACATGTCGGGGAGACT 6783
Qy 1660 TTGATGCAATTAATGCTGAACATTTTCACACGCGGACCCAGTATTTGAAACGACATTCG 1719
Db 6784 TTGATGCAATTAATGCTGAACATTTTCACAGAAAGTGAACCTGATTTGAGACGATATTCG 6843
Qy 1720 CGTGTGATTAAGCAAGCAAGCAAGCTATCGCCATTTGCGCGTTGATGATCTTGAAG 1779
Db 6844 COTGTTGACAAAGCAAGCAAGCAAGCTATGCGGTTAAACCGGCTGATGATCTTGAAG 6903
Qy 1780 ACTTATGCTGACCAACCGCTCTTAAATTTGATGAGCGCGCTTCCGCAATATCAT 1839
Db 6904 ACTGAGTGTGACCAACCACTACTCTGATTCGATTCGATGCGCTTTTGAAGAAATATCAT 6963
Qy 1840 CTGTGACCTACTACAGAAAGAGGTTTAAATTTGTCGATGATGAATCCGGTATGT 1899
Db 6964 CCACCATCTGCGCCAGGCTACCGCTTCAATTTGGGCGATGATGAATCCGAAATGT 7023
Qy 1900 TCTTAACGCTGTTTGTCAACACACTAGTCAATATCATGATTTGCTAGACAGATCTACGTG 1959
Db 7024 TCTTCAGCTCTTTGTCAACACAGTTTGAATGTCGTTATGCGCAGAGATTTGAGAG 7083
Qy 1960 AACGTTAACACAGCTCAGCGTCCGCGCTCTATGCGGAGATTAACATATGTCATGTG 2019
Db 7084 AGCGCTTAAACGTTCCAAATGTGACATTTATTCGGGAGACAACTTATATACCGGAG 7143
Qy 2020 TCGTCTCCGACACCTTGATGCGGAGAGATGGCCACTTGGCTGAACATGGAAGTAAAA 2079
Db 7144 TAGTATCTGACAAAGAAATGCTGAGAGTGTGCCACTTGGCTCAACATGAGGTTAAGA 7203
Qy 2080 TTATTTGATGCAATTAATGTTATCAAAAGCACTTCTGTGGGGATTTATCTGCTGTG 2139
Db 7204 TCATTTAGGAGTATATGCGGAGAGACCACTTACTTTGCGGTGATTCATCTTTCGAAG 7263
Qy 2140 ACCAGATTAACAGGACAGCTTGCAGAGTGCAGACCTCTTAAAGGCTTTTAAAGCTTG 2199
Db 7264 ATTCGTTACTCTCAACAGCGTGTGCGTGGGAGACCCCTTGAAGAGCTGTTTAAAGTTGG 7323
Qy 2200 GAAACCAATGCGAGTGAATATACCCAGACTGCGGACCGCGGCGCACTGCAATGATG 2259
Db 7324 GTAAACCGCTCCAGCCGACATGAGCAAGAGAAAGAGCCCTCTGCTAGATG 7383
Qy 2260 AAGCAATGCAATGGAACAGAAATTTGAATTAAGCAAGTATGTAAGGCGCGTGAATCCA 2319
Db 7384 AAACAAAGCGGTGTTTGAATGATGATTAACAGACACTTATGCAATGCGCGGCAATCC 7443
Qy 2320 GATACGAGATCATCTGCGAGCGCTGATCATGCTCTCTGTCACTGTTAGCCGAAAGCG 2379
Db 7444 GGTATGAGGTAGACAACTACACCTGTCTGCTGCACTTGAATCTTTTCCGACAGACA 7503
Qy 2380 TTAAGAACTTAAGAGCACTAAGAGGAGCCCAATACCCCTTAAGCGCTGAACCTTAATGAG 2439
Db 7504 AAAGGCAATTCAGACCATCAAGAGGGAATTAAGCATCTCTACGCTGCTCTTAATAGT 7563
Qy 2440 TGACGTAATGAG-----CAGCGCACTTACCCACCGGACGAGATGTTTCCA 2482
Db 7564 CAGCATTAATGATTTCAATCTGATTAATACCAACACACACACATGAATGAGGATTCCT 7623
Qy 2483 TACCTCAGCTGAACCTTTCAACAGTTTACCCCTAACAAATCCGATGCTTACCGAGATCCA 2542
Db 7624 TTAACATGCTGCGCGCGCGCCCTTCCAGCGCCCACTGCAATGTGAGCGCGCGAGAA 7683
Qy 2543 AACCTCTAGCGCGCGTGAAGCGCTTGGCGCCCGCTGCTGCTCAATGGAAT 2602
Db 7664 GAGGACAGCGCGCCCGATGCTGCGCGCAATGGAGTGGCTTCCAAATCCAGCAACTGA 7743
Qy 2603 CTTAGAGGTGATAGTCAACTTGACTTCAAAAC--GATCACTTAATCCGCGCGAG 2660

Db 7744 CCAAGCCGTCAGTGCCTAGTCATTTGAGACAGCAACTAGACCTCAAAACCCACAGCCAC 7803
Qy 2661 GTCCACCGCCAAAGAAAGAGAGTCTCTTAAGCCAAAACCTACTACGCTTAAAAAGA 2720
Db 7804 GCCCGCGCGCGCCAGAAAGAGCGGCCAAAGCAACCCAAACCCGAAAGAAACAA 7863
Qy 2721 AGAAGACGACGACCAAGGAGGAGCAAAAC---GCAAGCCTAAACGAGGAAACGACACGTA 2777
Db 7864 AAACACAGAGAGAGAGAGAGAGAGCACTCTGCAAAACCCAAACCCGAAAGAGACAGCTTA 7923
Qy 2778 TGTGTATGAAGTGTGAAGTCGAGACAAACATTTCCGATC---ATGCTGAACGCGCAAGTGA 2834
Db 7924 TGGCACTTAAGTGAAGCCGACAGACTGTCAGCTCAAAATAAGAGAGCGAAGATGCA 7983
Qy 2835 ATGCAATATGCTGCTGTGTGTGAGAGAAAGCTGATGAAACCTCCACGTTGAAGAAAA 2894
Db 7984 TCGGGACGACCTGGCCATGAGAAAGAAAGGTAATGAACCACTCAGCTGAAGAAACTTA 8043
Qy 2895 TTGATTAATGACCAATTAAGCGGCGGTGAATTTGAAGAAAGGTAGCATGTACGACTTGGAGT 2954
Db 8044 TTGACCACTCTGTCTATCAAAAGCTCAAAATTCACAAAGTCTCAGCATACGACATGAGT 8103
Qy 2955 ACGGCGACGTTCCCGAAGACATGAATTCGACACGCTGACAGTACACGACGCAAAACAC 3014
Db 8104 TCGCACAGTTCCCGTCAACATGAAGATGAGCGTTCACTTACCTACACGATGAACCCCTG 8163
Qy 3015 CGGGCTTCTAACACTGGGACCAACGCGCGAGTCCAGTATGAGAAATGGAGATTTACCGTAC 3074
Db 8164 AAGGGTTCTACACTGGGACACGAGGAGGTCAGTATGAGAGGAGATTTACCATCC 8223
Qy 3075 CGAAGAGAGTGGGCGGAAAGCGCACGACGGAAGCCGATCCTGGAACAACGAGGACAG 3134
Db 8224 CCCCGAGATGAGAGGACAGAGAGAGACAGTGTCTCTCGATTAATGATTAACCTCAGGCCG 8283
Qy 3135 TTGTGCTATTTCTTACAGAGTGCMAATGAGGCGACCGCTACGCGCTTTCACTGTGCA 3194
Db 8284 TTGTGCGATAGTCTCGGAGGGGCTGATGAGGGAAACAAACCGCCCTTTCCGTCGCA 8343
Qy 3195 CTTGGAACCAAGAGGGGTGACATTAAGGATACCCCGAAAGTTCTGAACCGTGT--- 3251
Db 8344 CCGTAATAGCAAGAGGAGCAATCAAGACACCCCGAAAGGACAAAGAGTGTCTG 8403
Qy 3252 -----CACTAGTTACAGCGCTATGCGCTTTGCAATGTCAAGTTCCATGCGCAAAAC 3305
Db 8404 CTGACCACTGTGTACCGGCACTGTGCTTTGGAAGCTGAGCTTTCCCATGCAATGCGCC 8463
Qy 3306 CACCGTGTCTATTCACCTGACGCGCAAGACCACTGACGCTGTCGAAGAGACGTCG 3365
Db 8464 CGCCCACTGTCTACCGCGCAACCATGCAAGCTCTGACATCTCTGAAGGAACGTGA 8523
Qy 3366 ACAATCCAAATTAAGACACGCTGCTGGAAGAGCTTTGAAT---GTCCATCAAGCGCGC 3422
Db 8524 ACCACGAGCCTACGACACCTGCTCAAGCCATTTGCGGTGCGGATCGTCGCGCAAA 8583
Qy 3423 CCAAAAGCAAGATTACCATGACTTCACTGACAGTCCCTACTCTGGGGTTCCCGCT 3482
Db 8584 GTAAAGAAAGGTCACTACGACTTTACTTGAACGCGCTGACTTTGGGACATGCTGT 8643
Qy 3483 ATTGCAACACTCAACGCGGTGTTCAAGCCCAATTAATTAAGAACGCTGTGGAGCAAT 3542
Db 8644 ACTGTCACTAATGTAACCGTGTCTTACCGGATTAAGATCGAGAGGTCTGGAGTGAAG 8703
Qy 3543 CTGATGATGATCGATTGAATCCAGGTCGCGCAAAATTTGGGCTCAATCAGGAGCA 3602
Db 8704 CGAGCAACACCACTACGACATACGACTTCGCGCAATTTGATGATACCAAAAGCGAG 8763
Qy 3603 CTGGGATGTCAACAAATTCGTTACATGTCTTTGACACAGACCATGACATCAAGAG 3662
Db 8764 CAGCAAGCTCAATTAAGTACCGTACATGTGCTGAGAGGATATACCTGTCAAAAGAG 8823
Qy 3663 ACAGATGAGAGAAATAGCTATCAGACATCTGACCTGCGCTGTCTTGGCCAAAG 3722
Db 8824 GCACATGATGATACATCAAGATCAGACCTCAGACCGGTGTAAGAGCTTAGCTCAAAAG 8883

Qy 3723 GGTACTTCTCTGTAGCTCAATGTCTCCAGGTGACAGTGTAACTGATACAGAGCG 3782
Db 8884 GATACCTTCTCTCGGAGAGTGTCTCCAGGGACAGCTTAAGCTTAGCATAGCGAGTA 8943
Qy 3783 GAGCATCTGAAATTCATGACCGGTGAGAAAAAGATCAGAGAGAAATTTGTGCTAGAG 3842
Db 8944 GCAACTAGCAAGTATGACAAATGAGCCCGCAAGATTAACCAAAATTTGTGGAGCGGG 9003
Qy 3843 AGGATTACTTGTTCCTCAACCGGTCCATGGAAGCTGTGAAGTCCACGTTTAAGTACT 3902
Db 9004 AAAAATATGACCTTACCTCCGTTTACGCTGAAGAGATTCCTTCAAGTGTACGACCTTC 9063
Qy 3903 TGAAGAGACGTTCGCGGGGTATACATAACATGACAGGCGGACCCACACGCTATTAAGT 3962
Db 9064 TGAAGAAACACCGCGCGCTACATCTATGACAGGCGGAGCCGACCTGCTATACAT 9123
Qy 3963 CTTATCTGAGAGAGCGTCAAGCGAAGTGTACTTAACCACTTCTGGCAAGAACGTCA 4022
Db 9124 CTTATCTGAGAGATCATCAAGGAAAGTTTACCGAAAGCCACCATCCGGGAAAGACATTA 9183
Qy 4023 CTTACGAATGTAAGTGTGGGACTACAGCACAGTATCTGAGCAGCGCAACGAAGTGA 4082
Db 9184 CTAAGAGTCAAGTGGGAGATTAACAAACCGGAACCGTTACGACCCGTACCGAAATCA 9243
Qy 4083 ACGGCTGACATAAAGCAAAACAGTGCATTGCTTACAAAGACGACCAAAAGAAATGGTCT 4142
Db 9244 CGGGTCAACCGCATATACAGTGTGCTCGCTTAAGAGTCACAAACGAAGTGGTCT 9303
Qy 4143 TCAACTGCGCGATCTTATTAAGCACACAGACCACTCAGTGAAGTAAATTGACATTC 4202
Db 9304 TCAACTGCGCGACTGATCAGACAGCGACGACCAAGCCCAAGGAAATTGATTTGC 9363
Qy 4203 CATTCGCTTGAACCGACAGTCTGCCCGGTTCGTTAGCTTACAGCCTTACAGTACGA 4262
Db 9364 CTTTCAAGCTGATCTCCAGTACCTGATGTCTCTGTTCACACGCGCCAAAGTATAC 9423
Qy 4263 AGTGTTCAAAGGATGACCTCCACCTGACTGCAATGAGCAGCAAACTTCTGACAGCA 4322
Db 9424 ACGGCTTAACATATGAGCTTCCATTAATGACACAGCACTTGACATTTGCTACACCA 9483
Qy 4323 GAAATTTGGGCTGCGAGACAGCAACAGACAGATTAAGAGGTCTTACATCCAGGA 4382
Db 9484 GGAATCTAAGGGGCAAAACCGGAACCAACATGATGATCATCGGAAACCGGTTAGAA 9543
Qy 4383 ATTTTCTGTGGGCGAGAAAGGCTGAGTACGATGAGGTAAACATGAACAGTCAAG 4442
Db 9544 ACTTCAACGTCGACCGAGATGCTGGAATACATATGGGCAATCAAGAACCAAGTAAAGG 9603
Qy 4443 TCTGGGCGCAGAGTGGGACACAGGCGACCAACATGATGAGCGCATGAGTCAATCC 4502
Db 9604 TCTATGCCAAGAGTCTGACACAGAGAACCTTCAAGATGCGCACAGAAATGATACAGC 9663
Qy 4503 ACTATTAATCACTGGCATCAAGTCTTACCTGATGTGTGTGTGCTGTCTTACTA 4562
Db 9664 ATTACTATATCGCATCTGCTGTGACACATCTTAAGCGCTGCACTAGCTGTGCGA 9723
Qy 4563 TCTGTGAGGCACTGCAATCATCAGCACTTGCATGCGCAAGCAAGAAAGAGACTGCTGA 4622
Db 9724 TGATGATTTGGGTAACTGTTCAGACATTAATGTCTTAAAGCGCGCGTGAAGTCTGA 9783
Qy 4623 CGGCATTCGCGCTGACACGGAACCAACCGGATACCAAGCATTAAGCGGTTTGTGCTGA 4682
Db 9784 CGGCATATGCTTGGCCCAATGCGGATTTCCAACTTGCTGTGCACTTTTGTGCTGTG 9843
Qy 4683 TTGCGCAACCAACGCTGAACATTTTGAAGAACTTTGAACCATCTGTGTTTAAACAC 4742
Db 9844 TTAGTGTGGCTAAAGTGAACATTTCAACGAGACCAATGATTAATTAATGTCGAAGCC 9903
Qy 4743 AACCGTTTCTGTGGCAACGTTGCAATTCCTCTGCGACGCGCTTGTATTTCTGTCCGT 4802
Db 9904 AGCGTTCTTGTGGTCCAGCTGTGTATCTCTGCGCGCTGTGTGTCTTAATGCGCT 9963

Oy	580	ACTGACGGTATGAAAGCGGAGCCGTAATATTTTCTCATCGGAAACAGGCCAAGTCACTTC	633
Db	5707	ACTGCTACACCGGGGTAGTGGGTGATCAATATTTTTCACGGAGACAGGCCCTTGGCACTTGC	5766
Oy	640	AACAGAAATCAAGTACGTCAATGTAAACTATACAGAAACATATATATGATCGGCCCGTCATG	699
Db	5767	AAAGAAATCCGTTCTGCAAGAACCAAGCTTACGAACCGAAGCTTGGAGGCAATGTCTCG	5826
Oy	700	AGAAATTTACGCCCCCGCCCTCGATCTTCGAAAGAGAGAAATGTTACAGAAAGATCG	759
Db	5827	AAAGATCTACGCCCCCGGTGCTCGACAGTCGAAAGAGAAACAGCTCAAACTCAGTAC	5886
Oy	760	AATTATGCGCCCTCGAAGAAATAGAACAGGTATCATTCAGAAAGTAGAAATATGA	819
Db	5887	AGATATGCCCAACCGAAGCCAAACAAAGAGTACCAAGTCTCGAAGATGAAACCAAG	5946
Oy	820	AAGCAATTAACGCGGAGGAGCTCATTTTGTGATTTGGGACATATCTATCATCAGAAAGTA	879
Db	5947	AAGCCATTAACCACTGAGGAGCTGTTTCAAGGGCTACGACTGATATATCTGCGACAG---	6003
Oy	880	ATCCTGTGAGTGTTCACAGAGTCATTTATCCTGTACCAATCTACTCTGTCACAGTAATTA	939
Db	6004	ATCAGCCAGAAATGCTATTAAGATCACTACCCGAAACCATGATTTCCAGAGTACAG	6063
Oy	940	ACAGGTTTACACTCTGCAAGGTCGCGGTTTAAACCTGCACACTAGTTATCCAGAGATTT	999
Db	6064	CGAAGCTACTGTGACCCCAAAAGTTTGTGTAGTGTGTTTGAACAACTATCTGATGAGATT	6123
Oy	1000	ACCTTACAGTACGCCAGTTATTTGTATTAACAGATGAATACAGTGCATCTTGACATGTTGG	1059
Db	6124	ACCCGACGGTAGCATCTTATCAGATCACCGAGAGTATGATGCTTATTTGATATGTGAG	6183
Oy	1060	ACGGGCACTCGTGTCTAGTACAGCCACTTTTGTTCGCGCTAACTGAGAAAGTCTCC	1119
Db	6184	ACGGGACAGTGGCTTGCTGATGATCTGCACTTTTGTGCCCGGCAAGCTTGAAGTTACC	6243
Oy	1120	CAAGAGACATAGCTATTTTGCAGCCAGATTAAGTCAAGCCGTCCATGCGCTATACGA	1179
Db	6244	CGAAAGACACGAGATATAGAGCCCAAACTCCGAGTGCAGTTTCATACAGATGCAGA	6303
Oy	1180	ATATCTTCAAAAATGTATTTGGTGAGCTACTMAAAGGAATTCGAAGCTTACCCAAATGC	1239
Db	6304	ACAGCTTCCAAAACGTGCTCATTTGCGGACTAAAGAAACCTGCACATCTCACAAATGC	6363
Oy	1240	GAGAAATTACTGTCTTAGATTTCGGCGGCAATTAATGTGATTTGTTTCAAGAAATACGAT	1299
Db	6364	GTGAATCTCCAAACACTGGACTGAGCATTCACAGTGAATGCTTTGAAAATATGAT	6423
Oy	1300	GCAATGATGAGTACTGGGATACCTTTCCGATTAACCTTATTCGGCTTAATCAAGAAACG	1358
Db	6424	GCAATGAGAGATATTTGGAGAGATTTGCCGAAACCAATTAGATATCACTACTGATTCG	6483
Oy	1360	TTACGCATATATGTGCAAAAGCTGAAAGGGCCGAAAGCAGCAGCATTTGTTGGAATACTC	1419
Db	6484	TTACCGCATACGTGGCCAGACTGGAAGGCCCTTAAGGCCGCCGCACTGTTCGAAAGACGC	6543
Oy	1420	ATATATCTAAAACGGTTGACAGAGATACCAATGATCAATTCGTCAATGATCTTAAAGAG	1479
Db	6544	ATATATTTGGTCCCATTTGCAAGAAAGCTATGAGATAGATTGCTCATGACATGAAAGAG	6603
Oy	1480	ATGTCAAAAGTTTACTCCCGGACAGAAACATACAGAGAGACGGCTTAAGGTGACAGTTATTC	1539
Db	6604	ACGTAAAGTTTACCTTGGACAGAAACACAGAAAGAAAGACCGAAAGTACAGTGAATAC	6663
Oy	1540	AGGCTGACAGATCCCTTGTCTACCGCTTACCTTTGCGGATCCATCGGAAATTAGTCCGTA	1599
Db	6664	AAGCCGAGAAACCCCTGGGACCGCTTACTATGCGGATCCACCGGAGTTAATGTCGCA	6723
Oy	1600	GACTGAAATGCGGTGCTTCTGCCAAATATCATCTCTTTGGACATGTTCAGCGGAAGTT	1658
Db	6724	GGCTTACAGCCGTTTGTCTACCCACATTCACACCTCTTTGACATATGTCGCGGAGACT	6783

Qy	1660	TTGATGCGATATATGCTGGAACATTTCCACACGCGCGACCCAGTTTGGAAACGGACATCG	1720
Db	6784	TTGAATGCAATATATGACGAACACTTCAAGCAAGGTGACCCGGTACTGAGACGGATATCG	6843
Qy	1720	CGTGGTTTGAATAAAGCGAAGACGACGCTATCGCATTTTCGGCGCTTGATGATCCTTAGG	1779
Db	6844	CCTGGTTTGAACAAAGCCMAACGACGCTATGCGGTTTAAACCGGCTGATGATCTTTGGAG	6903
Qy	1780	ACTTAGGTGTGACCAACCGCTCTTGAATTTGATAGAGCGCGCTTCGGCAATATCAAT	1839
Db	6904	ACCTGGGTGTGACCAACCACTTACCTGACCTTGATCGAGTGGCGCTTTGGAGAAATATCAT	6963
Qy	1840	CTGTGCACTTACCTTACAGGAACGAGGTTTAAATTTGGTGCATATGATAAATCCCGTATGT	1899
Db	6964	CACTCCATCTCTCCACGGGTACCCGTTTCAAAATTCGGGCGCATATGAAATCCGGAATGT	7023
Qy	1900	TCTTAAGCGTGTGTTGTCAACACACTAGTCAATATCATATGATTTGCTAGCAGAGTACTACGTG	1959
Db	7024	TCTTACGCTCTTGTGTCAACACAGTTCTGAAATGTCTTATGTGCGACGAGATATTTGAGG	7083
Qy	1960	AACGGTTAACGACGTCAAGCGTGCAGCGCTCTATCGGCGACGATPAACAATAGTCATGGTG	2019
Db	7084	AGCGGCTTAAACGTCGAATATGTGACAGATTTATCGGAGAGCAACATTTATCACGGAG	7143
Qy	2020	TGCTGTCCGACACTTGTATGGCGGAGAGATGCCGCACTTGGCTGAACATGGAATMAAA	2079
Db	7144	TAGTATCTGCAAGAAAGAAATGGCTAGAGGTGTGCACCTGGCTCAACATGGAAGGTTAAGA	7203
Qy	2080	TTATTTGATGACGTTATTTGGTATCAAAAGACCCCTACTCTGAGGGGGAATTTATCCTGGTGG	2139
Db	7204	TCATTTGACGAGTCAATCGGCGAGAGACCACTTACTTCTGGGTTGATTCATCTTGCAGG	7263
Qy	2140	ACCAATPAACGAGCACAGCTGCAGAGTCCGAGACCTCTMAAAAGCTTTTAAAGCTTG	2199
Db	7264	ATTTGGTTACTCCACACAGCGTGTGCGGTTGGGAGACCCCTTGAAGAGGCTGTTTAAATTTGG	7323
Qy	2200	GAAACCATTTGCCAGTGCATATATACCAAGCTGCGACCGCGCGCGGCACTGCATGATG	2259
Db	7324	GTAACCCCTCCACGCGCAGCATATGACAGACGAAAGACAGAAAGACGCGCTCTGTAGATG	7383
Qy	2260	AAGCAATGCGATGGAACGAATTTGSAATTTACGACGAGTTGATGTAAGGCGGTAGAAATCCA	2319
Db	7384	AAACAAAGCGGTGTTTATGATGATATTAACAGACACTTTAGCACTGGCCGTGGCAATC	7443
Qy	2320	GATACGATCATATCTGGCAGGCGCTGATCATCAAGTCTCTGTCCACGTTAGCCGAAACCG	2379
Db	7444	GGTATGAGATGACAAACATCAACACTGTCTGTGGCACTTGAGAACTTTTGGCAGAGCA	7503
Qy	2380	TTAAGACTTCAAGAGCATTAAGGAGAGCCCAATCACCTTATACGCTGACCTTAAATAGG	2439
Db	7504	AAAGGACATTTCAAGCCATCAGAGGGGGAATTAAGCACTTCTGAGGTGTCTTAAATAGT	7563
Qy	2440	TGACGTATTAAC-----CAGCGACCTAACCCACGGGAGAAATTTTCCA	2482
Db	7564	CAGCATATGATTAATTTCACTGATCTPAATACCAACACACACCACTGATATAGAGATTTCT	7623
Qy	2483	TACCTCAGCTGAACCTTCCACACGATTTTACCTTCAAAATCCGATGAGCTTACCGAGATCCA	2542
Db	7624	TTAATATCTCGGCGCGCGCCCTTTCCAGGCCCACTGCCCATGTGGAGGCGCGGAGAA	7683
Qy	2543	AAACCTCTTAGGCGCGCTGAGAGCGCTTTCGCCCCCGCTGCTCTCAATGGAAGAT	2602
Db	7684	GGAGGCAAGGCGCGCCGATGCTGCTCCGCAATGGGCTGCTTCCCAATCCAGCAACTGA	7743
Qy	2603	CTTAGGAGTGCATATGATCACTTGAACCTTTCAAACAC--GATCACTAATTCGCGCGCAG	2666
Db	7744	CCACAGCGGTAGTGCCTTATCTATTGACAGGCAACTAGACTTCGAACCCCAAGCCAC	7803
Qy	2661	GTTCACCGCCAAAGAAAGAAAGATGTCTTAAAGCCAAACCTACTCAGCTTAAAGA	2720
Db	7804	GCCGCGCGCGCGCCAGAAAGACAGGCGCCAAAGCAACACCGAAGCCGAAGAAACCA	7863
Qy	2721	AGAACGACGAAGCAAGAGCAAGAAC--GCAAGCTTAAACAGGGAACGACACAGTA	2777

Db	7864	AAACACGAGAAAGAAAGAAAGAACCACTTCGAAAACCCAAACCCGGAAAGAAAGACGCTA	79233
OY	2778	TGTGTATGAAGTTGGAGTCGGAACAAGACATTTCCGATC---ATGTGAACGGCCAAAGTGA	28344
Db	7924	TGGCACTTAAGTTGGAGGCCGACAGACTGTTCGACGTCAAAAATATGAGACCGAGATGTCA	79833
OY	2835	ATGGAATATGCTTCGCTTGTCCGAGGAAGGCTGATGAAAACACTCCACTTGTAGAGAAAAA	28944
Db	7984	TCGGGACAGCACTGGCCCATGGAAGGAAAGTAATGAAAACAATCCACTCACTGAAAGAACTA	80433
OY	2895	TTGATTAATGACCAATTTAGCGGCCGCTGAATTTGAAGAAGGCTAAGCATGTAGCACTTGAAGT	29544
Db	8044	TTGACACACCTCTGTGTATCAAAAGCTCAATTTCAACCAAGTTCGTACAGATTAACATATGAGT	81033
OY	2955	ACGGGACGTTCCCAAGAACATGAAATCAGACACGCTGCAGTACACACGCAACAAACAC	30144
Db	8104	TCGCACAGTTCGGGTCAACATGAGAAGTGAAGGGCTTCACTTAACACAGTAAACACCTGT	81633
OY	3015	CGGGCTTTACAATCGGCACCAACGGCCGACATCCAGTATGAAATGGAGATTTACCGTAC	30744
Db	8164	AAGGGTTCTCAACTGGGCACCAACGAGGGCGGTGCAGTATAATGTGAGGCGAGATTTACATCC	82233
OY	3075	CGAGAGAGTGGGGGGGAAAGGCGCACGCGGAACCGATCTCGGACAAACAGAGGCGAG	31344
Db	8224	CCCGGGAGTGAAGAGCAGAGAGACATGTGTCGTCGATTAATGATTAATCAAGCGCGGG	82833
OY	3135	TTGTGTCTATTGTTCTTAAGAGGTGCMAATGAGGCAACCGTACCGCGCTTCAATGTGCA	31944
Db	8284	TTGTGTGCGATAGTCTCTCGGAGGGCGCTATGAGGAAACAAGAACCGCCCTTTCGGTGTCA	83433
OY	3195	CTTGGAACCAAGAAAGGGGTGAACCATTAAGGATACCCCGAAGGTTCTGAAACCGTGT---	32513
Db	8344	CCTGGAATAGCAAAAGGGAAGCAATCAAGAACACCACCGAAGGGAAGAAAGTGTCTGT	84033
OY	3252	-----CACTAGTTTACAGCGCTAAGCGCTAAGCGGCTTTGGAATGTCAAGTCCCATATGCGAACAC	33053
Db	8404	CTGCACCACTGTCTACCGGCAATGTGCTTTGGAAAGTGAAGCTTCCCATATGACATCGCC	84633
OY	3306	CACCCGTGTGTATTCACTGACGCGCAGAACAAACACTGACGTGCTCGAAGAGAACGTGG	33653
Db	8464	CGCCCACTGTTCACACCCCGGCAACCATCAAGACTCTGGAATCTCTGAAAGAAACGTGA	85233
OY	3366	ACAAATCCAAATTAACGACGCTGCTGGAAGAAAGTCTTGAAAT---GTCCATCACGCGGC	34223
Db	8524	ACCACGAGGCTTACGACACCTGTCTCAACGCAATTTGCGGTGCGGATCGTCCGGCGAA	85833
OY	3423	CCAAACGAAGCATTAACGATGACTTCAACACTGACCACTGACCAAGTCCCTACTGGGGTTTCGCCGT	34823
Db	8584	GTAAGAAGAGGTCACTGACGACATTTTACTTGAACGACGCCCTACTTGGGCAACATGCTCGT	86433
OY	3483	ATTGCGAGACCTCAACGCGCGTGTTCACGCCCAATTAATAATTGAGAAGCGTGGGACGAT	35423
Db	8644	ACTGTACCACTACTGGAACCGTGTCTTTAGCCCGAATTTAAGATGAGCAGAGTCTGGGAATGAG	87033
OY	3543	CTGATGATGATCGATTGAATTCAGAGTCTCGGACCAATTCGGCTACATTCAGCAGGCA	36023
Db	8704	CGGACGACAAACCACTATGACATPACGAACTTCGCGCCAGTTTGATAGGACCAAAAGCGAG	87633
OY	3603	CTGGGGAATGACAAATTCGGTTACATGTCTTTGACCAACGACCATGACATCAAGAG	36623
Db	8764	CAGCAAGCTCAAAATAGTACGCTACATAGTGCCTGACAGAGATCATCTGTCAAAAGAG	88233
OY	3663	ACAGATGAGAAAAATAGCTATCAGCAATCTGGACCTGCGCGTCTTTGGCCACAAG	37223
Db	8824	GCAACATGATGACATCAAGATCAACACCTCAGACCGCTGTAGAAAGCTTAGCTTACAAAG	88833
OY	3723	GGTACTTCTGTAGCTCAATGTCTCTCCAGGTGACAGTGAACCGTCAAGTATCAACGACG	37823
Db	8884	GATACTTCTCTCTCGCAAGTGTCTCTCCAGGGGACAGGTAAACGGTTTAGCATTAACGAGTA	89433
OY	3783	GAGCATCTGAAGATTCATGCAACCGTGGAGAAAAAAGATCAGAGGAAAGTTTGTCCGTAGAG	38423

Db	8944	GCAACTGACGACGTCAATGCAACATGCGCCGCAAGATTAACCAAAATTTCTGGACGGG	9003
OY	3843	AGGAGTACTTGTTCACACCCGTCATGAAAAGCTGTAAAGTGCACGTTTACATCACT	3902
Db	9004	AAAAATATGACCTAACCTCCCGTTACAGGTAAAGAAATTCCTTTCACAGTGTACGACCGTC	9063
OY	3903	TGAAGGAGACGTCTGCCCGGTTACATTAACATGCAACGCGCCACACACGCTATTAACT	3962
Db	9064	TGAAAGAAACACCGCCGGCTTACATCACTATGCAAGCGCGGGAACGCAATGCTTAATCAAT	9123
OY	3963	CCATCTGAGGAAAGCGTCAGGCGAAAGTGAACAATTAAACCACTTTCTGGCAAGACGTA	4022
Db	9124	CCATATCTGGAGGAATCATCAGGAAAAGTTTACCGGAAGCACCATTCGGGAAGACATTA	9183
OY	4023	CCTACGATATTAAGTGTGGCACTAAGCAACAGTATCTGTAGACACGCGAAACGAATGA	4082
Db	9184	CGTACGAGTGCAAGTGGCGGATTAACAAGACGGAAACCGTTACGACCCGTAACGAAATCA	9243
OY	4083	ACGGCTGCATTAAGCAAAAACATGTCAATTGTCTTCAAGAGCGAACCAAGAAATGGGTCT	4142
Db	9244	CGGGCTCACCGCCATCAACAGATGGGTCTGCTTAAGAGCGACCAAGAAAGTGGGTCT	9303
OY	4143	TCAACTGCGCGAATCTTATTAAGGCAACAAGACCACTCAGTGAAGTAATTTGCACATTC	4202
Db	9304	TCAACTGCGCGAATCGATCAGACAGCGCACACACGGCCCAAGGAAATTTGCATTTTC	9363
OY	4203	CATTCCGCTTGACACCGACAGTCTGCGCGGTTCCGTTAAGCTCAACACGCTACAGTCA	4262
Db	9364	CTTTCAAGCGTATCCGAGTACCTTGATGGTCCCTTGTTCGCCACGCGCGGAAGTATGATAC	9423
OY	4263	AGTGTTCAAAGGATCACTCTTCAACCTGACATGCAATGCGAACACATTTGCTGACACGA	4322
Db	9424	ACGGCTTTAAACATCATGACCTCAATTAGACACACACATGCAATTTGCTTACACCA	9483
OY	4323	GAAATTTGGGGCTGCGACGACGCAACAGCGAATGAAATTCAGGCTTCAATCCAGCA	4382
Db	9484	GGAACCTAGGGGCAAAACCGGAAACCAACACTGAATGATTCATCGGAAAACACGGTTAGAA	9543
OY	4383	ATTTTCTGTGGGGCGAAGAGGCTGTGAAGTACCTATGGGGTTAAACATGAAACAGTCA	4442
Db	9544	ACTTCACCGTGCACCGAGATGGCTGGAAATACATATGGGGCAATCAGAACAGTAAAGG	9603
OY	4443	TTGGGGCCCAAGAGTGGGACCAAGGCGAACCCACATGGATGGCCGACGTGATCATATCC	4502
Db	9604	TCATGCCCAAGAGTGTGCACACGAGAGACCCCTCAGATGGGCCACACGAAATATGTA	9663
OY	4503	ACTATTTCATCGGCATCCAGTCTTAACATGTCTATTGCTGTGTGGTGTCCCTTCTTGTCTA	4562
Db	9664	ATTACTTCATTCGCAATCCCTGTGTACACCATCTTTACCCGTCGATCAAGCTGTGGCGCA	9723
OY	4563	TCTGTGATGAGCACTGCATCATCAGACAGCTTGCACTCGCCAAAGCAAGACAGTGCCTGA	4622
Db	9724	TGATGATTTGGCGTAACTGTGTGACGCAATTATGTGCTGTAAAGCGCGCTGAGTGCCTGA	9783
OY	4623	CGCCATACGGGCTTGCACCGAAACGCAACGGTATCCCAACAGCATTTAGCGATTTTGTGCTGA	4682
Db	9784	CGCCATATGCGCTGTGGCCCCCAAAATGCGGTATTCMACTTCGATGGCACTTTTGTGTGTG	9843
OY	4683	TTGGGCGCAACCAACGCTGAAACATTTGGAGAAACTTTGAACCATCTGTGTGTTTAACACC	4742
Db	9844	TTAGTGTGGGTAAATGCTGAAACATTCACCGAAGCACTGATGATTAATTATGTTGTAACAGCC	9903
OY	4743	AACGTTTCTGTGGGCAAGTTGTGCAATTCCTGTGACAGCGCTGTGTTATTCTGTTCGGCT	4802
Db	9904	AGCGCTTCTTGTGGGTTCAGCTGTGTATACCTTGTGGCCGCTGTGTCGTTCTTAATGCGCT	9963
OY	4803	GCTTTTCATGCTCATGCTTTTTTTATTTGATTTGACAGCGCTGTGCTGTGGGAAAGTATACG	4862
Db	9964	GTTGCTCATGCTGCGTCCCTTTTATTTGATGGTTGCGCGCCCTTACCTGCGGAAGTATGACG	10023
OY	4863	CCTTTCAAACTGGACCACTGTGCGCAAAATGTTTCGGGGATCCCGTATTAAGCGCTTGTGCG	4922
Db	10024	CTTACGAAACATGGACCACTGTGTTCAAAATGTGCAACAGTATCCGTAATAGGCACTTGTGTG	10083

QY 4923 AACGCGAGTTAGCGCACTTAACCTGAGATCACGGTGTCTCATCGGAATTAAAC 4982
 DB 10084 AAGAGGAGGGTACGCCCGCTCAATTTGGAGATTACTGTATGTCTCTGGAGGTTTGC 10143
 QY 4983 CTTCACATAACAGAGTACGTGACCTGCAAAATTCACACAGTATTCCTTACACACAG 5042
 DB 10144 CTTCACATAACAGAGTACGTGACCTGCAAAATTCACACAGTATTCCTTACACACAG 10203
 QY 5043 TTAAATCTGCGGGTCTCTGAGTGCAGAGATCTCAAGAGCGGATTTACATGCCGCG 5102
 DB 10204 TCAGATCTGCGGGTCTCTGAGTGCAGAGATCTCAAGAGCGGATTTACATGCCGCG 10263
 QY 5103 TTTTTCGCGGTGTACCTCTTCACTGTGGAGGCGCAATGCTTCTGTACAGTGA 5162
 DB 10264 TCTTTGAGGGGTATACCTCTTCACTGTGGAGGCGCAATGCTTCTGTACAGTGA 10323
 QY 5163 ACACACAATAGTAGGCGGTACGTGAGTTGCTCCAGACTGTACATATAGATCACGCG 5222
 DB 10324 ACAGCCAGATAGTAGGCGGTACGTGAGTTGCTCCAGACTGTACATATAGATCACGCG 10383
 QY 5223 TCGCACTAAAAGTTACACAGCTGTCTGAAAAGTGGCGCTGTATAGTATACGGCA 5282
 DB 10384 AGGGATTAAGTGCATATGCGCGAGTAAAGTAGAGTGGTATAGTATACGGCA 10443
 QY 5283 CCACCGCGCACTGTGATAGTTGTCAATGCGGTACGCGAGGTTCTTCAAGGACCTGA 5342
 DB 10444 CTACCACTTTTCTAGATGTGTACGTGAACGAGTACACCGAGAACTCTTAAAGCTGA 10503
 QY 5343 AGGTATAGCAGGCGCGATACCGCTTTTTCACCTTTGACCATTAAGTGTCTATCA 5402
 DB 10504 AAGTATAGTGTGACCAATTTTACGATTTTACACCTTGATGATCAAGGTCGTATCA 10563
 QY 5403 GAAAGGCTTGTTTTACATCACTACGATCTTCTGATGTAGAGTATGAAACAGAGGCT 5462
 DB 10564 ATCCGCGCTGTGTACAACTATGACTTTTTCGGAATGCGAGATGAAACAGAGGCT 10623
 QY 5463 TCGCGATATCAAGCATCTCTGCTTGTATGCTACAGCATAGTAGGCGCGCATGATAC 5522
 DB 10624 TTGAGACATTCAGGCTACCTCTTGTACTACAAAGACTCATGCCGACACATATTA 10683
 QY 5523 GGTGTGAAAGCTTCTGTCAAGAACATCACTGCTCCCTACACCCCAAGCATATCAGGT 5582
 DB 10684 GGCTACTCAAGCTTCCGCAAGAACGTGATGCTCCGTAACAGCAGGCGCCATCTGGAT 10743
 QY 5583 ATGAAATGTGAAGAACAACTACAGACGACCCCTGCAAGAAACAGCACTTTGGATTA 5642
 DB 10744 TCGAGATGTGAAAAAACTACAGCGCGCCCACTGCAAGAAACGCGCTTTTGGTGCA 10803
 QY 5643 AAATTTGAAGTGAAGCTCTGCGAGGTACTAATGCTTACGGGCAATCTCTATCTGA 5702
 DB 10804 AGATTTGAGTCAATTCGCTTTGAGCGGTGACTGTCTATACGGAACTTCTCAATTTCTA 10863
 QY 5703 TTGACATCTCGATGACACTTTTGTAGATCATCAAGATCAACAAATTTTGAAGTTA 5762
 DB 10864 TTGACATCTCGATGACACTTTTGTAGATCATCAAGATCAACAAATTTTGAAGTTA 10923
 QY 5763 GCTGACATGACGACTGTCTATTTTGTGACAGCTTTGTGTCTTAACTATACGT 5822
 DB 10924 AATGTGATGTACGTGACGCTTATTTGAGGAGCTTGTGAGGAGTGTCTACCTGTGAGT 10983
 QY 5823 ACAAGCTGACAGGAGGAGCAATTTGCAAGTTCACTCCACCTCCACGACAGCTGTTTGA 5882
 DB 10984 ATGTATCTGACCGGAGAGGACAAATGCTGTATCACTTCCATTTGACAGCAAGCACTTCC 11043
 QY 5883 AGAAGACGACCAACATGTGACTGCGGTAGGACAGATTAACATATTTTGAAGTGA 5942
 DB 11044 AAGAGTGAAGTTCATGTCTCGAGAAAGAGCGGTGACAGTATCACTTACGACCGGCA 11103
 QY 5943 GCCCAAGCAAAATTTTATGTTTGTATGCTATGCGGCAAGAGTCCACTGTCAATGTGAT 6002
 DB 11104 GCCCAAGCGCAACTTCAATGTATGCTGTGTGTGAAGAGCAACATGCAATGTAGAT 11163

QY 6003 GTAAACCAACCGCGCAACCATATATGAGAACCAATAAATGCAACCAAAATTCAGG 6062
 DB 11164 GCAAAACCAACCGCGATCATATGTGTAGACACCCCGACCAAAATGACCAAGAAATTCAG 11223
 QY 6063 CGGCACTTCCAAACATCTTTGGAACCTGCTGCTGTGACATCTTTGGGAGCATATCC 6122
 DB 11224 CCGCATCTCAAAAACCTTCAATGAGTGTGCTGTTTGGCCCTTTTGGGCGGCTGTGC 11283
 QY 6123 TCAATGTTAGAGCTTATATGTGTGTGCTGACAGCTATGCTTATTAACACAGTATGAT 6182
 DB 11284 TATTAATTAATAGACTTATATATTTTCTTGACATGATGCTGACTACAGCAAGAT 11343
 QY 6183 GACTGACCGGCACTGA-CATAGCGTAAACTGATATCTTCCGAGAAAGTGTG 6241
 DB 11344 GACCGTACGCCCAATGACCCGACGCAAACTGATGTACTTCCGAGGAAGTATGT 11403
 QY 6242 GCATATGCCAGCGCGCT-----TGACACT 6288
 DB 11404 GCATATGCAATCAGGCTGTATATATATGATCCCGCTTACCGCGGCAATATAGCAAC 11463
 QY 6289 AAAATCTGATATTTTCCGAGGAAGCAGTGCATATGCTGTGACAGT 6318
 DB 11464 AAAATCTGATATTTCCGAGGAAGCAGTGCATATGCTGTGACAGT 11513

RESULT 11
 146902
 LOCUS 146902 11663 bp DNA linear PAT 07-OCT-1997
 DEFINITION Sequence 1 from patent US 5639650.
 ACCESSION 146902
 VERSION 146902.1 GI:2470867
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 11663)
 AUTHORS Johnston, R.E., Davis, N.L. and Simpson, D.A.
 TITLE CDNA clone for South African Arbovirus No. 86
 JOURNAL Patent: US 5639650-A 1 17-JUN-1997;
 FEATURES
 source 1..11663 /organism="unknown"
 BASE COUNT 3291 a 3105 c 2907 g 2360 t
 ORIGIN

Query Match 36.9%; Score 2396; DB 6; Length 11663;
 Best Local Similarity 64.6%; Pred. No. 0;
 Matches 3751; Conservative 0; Mismatches 1985; Indels 74; Gaps 9;

QY 580 ACTGACGTTAAGACGGGAGCGTATATTTTCTCATCGAAACAGGCCAAGTCACTTC 639
 DB 5707 ACTGTTAACCAGGGGTAGTGTGATATTTTTCAGACAGACAGGCCCTGGGCACTTGC 5766
 QY 640 AACGAAATCAGTACGTCATATGTAATCAAGAACTATATTTGATCGGGCCTCATG 699
 DB 5767 AAAAGAGTCCGTTCTGCAAGACAGCTTACAGAACCGACTTGTGAGACGCAATTTCTGG 5826
 QY 700 AGAAGTTATACGCCCCCGCTGATCTCGAAAGAGGAAATTTTACAGAAAGAACTGC 759
 DB 5827 AAAAGATCTACGCCCGCTGCTGACAGTGCAGAAAGGAAACAGCTCAAACTCGTACC 5886
 QY 760 AATATGCGCTCTGAAGGAATATGAAAGAGGTATCAATCAAGAAAGTGAATATGA 819
 DB 5887 AGATGATGCCACCGAGGACCAACAAAGAGTACAGTCTCGAAAAAGTGAAGAAACAGA 5946
 QY 820 AAGCAATTAACGCGAGCGACTATTTCTGATTTGGGACATATCTATCATCAGAGTGA 879
 DB 5947 AAGCATTAACCATGACGACTGCTTTCAGGGCTACAGATGATTAATCTTGCACAG--- 6003
 QY 880 ATCTGTGAGTGTTAACAGATCAATATATCTGTGACCAATCTACTGTCACAGGTATTA 939
 DB 6004 ATCAGCAGATGCTATTAAGATCACTTACCGAAACATGATATTCAGCAGGTGTACAG 6063

Qy 940 ACAGTTTACATCTGACAGAGTCCGGGTTAAACGTGCACTTATGTTATCCAGAGAAATT 999
Db 6064 CGAACTACTCTGACCCAAAGTTTGCTAGTGTGTTATGTAACAACATCTGCAATGAGAAATT 6123
Qy 1000 ACCCTACGTAAGCCAGTTATGTTATTAACAGATGAATACGATGCGTATCTTGACATGTGG 1059
Db 6124 ACCCGACGGTAGCATCTTATCAGATCACCAGACGATGATGCTTACTTGGATATGTAG 6183
Qy 1060 ACGGCGCATCGTGTCTAGTATACGCACTTTTGTCCCGCTTAACCTGAGAACCTACC 1119
Db 6184 ACGGACGCTGCTGCTTCCCTAGTACTGAACTTTTGTCCCGCTTAACCTGAGAACCTACC 6243
Qy 1120 CAAGAACATAGCTATTTGACGAGCATAGATTCAGCCGCTCCATCGCTATACAGA 1179
Db 6244 CGAAAAGACGAGATATAGAGCCCAAAACATCCGAGTGGCTTCATACAGATGAGCA 6303
Qy 1180 ATACATTCAAAATGTATGCTGCTGCACTTAAAGAAATGCAAGCTTACCAATGC 1239
Db 6304 ACAGCTTGCAAAACGTGCTCATTTGCGGACTAAAGAACTGCACACCAAAATGC 6363
Qy 1240 GAGATTACCTGTCTTATGATTCGGGCGCATTTAATGTTGTTTCAAGAAATCGCAT 1299
Db 6364 GTGAATCTCCAACTGACTGACCTGACGACATTCACGTTGAATGCTTGGAAAATATGAT 6423
Qy 1300 GCAATGATGATGATGAGTACCTTTCCGCAATACCTTATCGCTAATACTACAGAAAG 1359
Db 6424 GCAATGATGATGATGAGTACCTTTCCGCAATACCTTATCGCTAATACTACAGAAAG 6483
Qy 1360 TTAAGCAATATGTGCAAAAGCTGAAAGGCGCAAGACAGCATTTGTTTGGCAATATCTC 1419
Db 6484 TTACCGCATATGTCGCGCAGACTGAAAGGCGCTTAAAGGCGCGCATGTTTGGCAAGACG 6543
Qy 1420 AATAATCTAAAACCGTTGACAGAGATACCAATGATCAATTTGTCATGATCTTAAAGAG 1479
Db 6544 AATAATTTGCTCCATTTGCAAGAAAGTCCCTATGATGATGATGCTCATGATGATGAAAG 6603
Qy 1480 ATGTCAAAATTAATCTCCGCAAGAAACATACAGAGAGCGGCTTAAGGTGACGTTATTC 1539
Db 6604 ACGTGAAGATTAATCACTGCGCAAGAAACATACAGAGAGCGGCTTAAGGTGACGTTATTC 6663
Qy 1540 AGGCTGAGATCCCTTGTCTACCTTACCTTATCGCTTATCGGATCCATGCGGATTTAGTCC 1599
Db 6664 AAGCGCGCAAGCCCTGCGCAAGCTTATCTATGCGGATCCATGCGGATTTAGTCC 6723
Qy 1600 GATGTAATGCGGTGTTTGTCCCAATATCAATCTCTTGTGACATGTCAGCGGAATTT 1659
Db 6724 GGTCTTACGCGGTTTGTCTACCAATTCACACCTCTTGTGACATGTCAGCGGAATTT 6783
Qy 1660 TTGATGCGATTAATGCTGAACATTTCCACACGCGGACCCAGTATTTGAAAACGAGATGC 1719
Db 6784 TTGATGCGATTAATGCTGAACATTTCCACACGCGGACCCAGTATTTGAAAACGAGATGC 6843
Qy 1720 CGTGTGTTGATTAAGCGAAGACGACGCTATCGCCATTTGCGGCTTGTGATGATCTTTGAG 1779
Db 6844 CCTGTGTGACAAAAGCCAAAGACGACGCTATGCGGCTTGTGATGATCTTTGAG 6903
Qy 1780 ACTTATGCTGATGACCAACCGCTCTTATGATTTGATGAGCGCGCTTGGCAATATCAT 1839
Db 6904 ACTTATGCTGATGACCAACCGCTCTTATGATTTGATGAGCGCGCTTGGCAATATCAT 6963
Qy 1840 CTGTGACCTTACCTACAGAGAGAGTTTAAATTTGATGCGCATGATGAAATCCGATATGT 1899
Db 6964 CCACCATCTGCTCCACGAGTACCGCTTCAAATTTGGGCGCATGATGAAATCCGATATGT 7023
Qy 1900 TCTTAAAGCTGTTGTCAACACATGATCAATATCATGATTTGATGAGAGTACTACGTG 1959
Db 7024 TCTTAAAGCTGTTGTGTCAACACATGATGATGATGATGATGATGATGATGATGATGATG 7083
Qy 1960 AAGGTTTACCAAGCTGACGCGGCTGATGAGGAGATTAACATAGTGCATGATG 2019
Db 7084 AAGGTTTACCAAGCTGACGCGGCTGATGAGGAGATTAACATAGTGCATGATG 7143
Qy 2020 TCGTCTCGACACTTATGAGCGAGAGATGCGCACTTTGGCTGAACATGAGAAATGAAAA 2079

Db 7144 TAGTATCTGACAAAAGAAATGCTGAGAGTGTGCACTGCTGCTCAACATGAGGTTAAGA 7203
Qy 2080 TTATTTGATGATGATTTGATTTCAAGACCTTACTTGTGAGGAGATTTATCTGTGG 2139
Db 7204 TCATTTGATGATGATTTGATTTCAAGACCTTACTTGTGAGGAGATTTATCTGTGG 7263
Qy 2140 ACCGATTAACAGGACGAGCTGAGAGTGTGAGAGCCCTTAAAGGCTTTTAAAGCTTG 2199
Db 7264 ATTGATTTACTTCAACAGCTGTGCGTGGGAGCCCTTAAAGGCTTTTAAAGCTTG 7323
Qy 2200 GAAAACCATTTGCGATGATATCCCAAGCTGAGACCGCGCGGCTGATGATGATG 2259
Db 7324 GTAAACCGCTTCCAGCGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7383
Qy 2260 AAGCATGCTGATGAGAGAGATTTGAGATTTACGAGAGATTTAGTGAAGCGCTGATGAT 2319
Db 7384 AAACAAAGCGTGTGTTATGATGATTTACGAGACCTTATGAGTGTGCGTGTGCAATC 7443
Qy 2320 GATACGATGATATGCGAGCGCTGATCATGATCTGTGCTGCACTTACCGTAAAGCG 2379
Db 7444 GGTATGAGTATGACAAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7503
Qy 2380 TTAAGAACTTCAAGAGCATTAAGAGAGAGCCCAATCACCTTACGCTGATGATGATG 2439
Db 7504 AAAGAGATTTCAAGCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7563
Qy 2440 TGACGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2482
Db 7564 CAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7623
Qy 2483 TACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2542
Db 7624 TTAACATGCTGCG 7683
Qy 2543 AACCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2602
Db 7684 GAG 7743
Qy 2603 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2660
Db 7744 CCACAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7803
Qy 2661 GTCCAGCGCAAG 2720
Db 7804 GCG 7863
Qy 2721 AGAAGCAGCAAGCAAG 2777
Db 7864 AAACAG 7923
Qy 2778 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2834
Db 7924 TGGCATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7983
Qy 2835 ATGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2894
Db 7984 TCGGACAGCACTGCGAG 8043
Qy 2895 TTGATTAATGAGCAATTTGAGCGCGCTGAAATTTGAAGAGGCTGATGATGATGATGAT 2954
Db 8044 TTGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8103
Qy 2955 ACGGCGAGCTTCCCAAGACATGAATCAAGACGCTGAGTACCAACGAGCAAAACCAAC 3014
Db 8104 TCGCACAGTTGCGCGTCAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8163
Qy 3015 CGGCTTCTTCAATTTGCAACCAAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3074
Db 8164 AAGGCTTCTTCAATTTGCAACCAAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8223
Qy 3075 CGAGAGAGTGTGCGGAG 3134

D8 8224 CCCGCGAGTAGAGAGAGAGAGACAGTGGTCGTCGCAATTATGAACTCAGCGCGG 8283
QY 3135 TTGGGCTATTGTTCTAGAGAGTGCMAATGAGGCAACGCGTACCGCGCTTCACTGGCA 3194
D8 8284 TTGTCGGATGATGCTCGAGAGGGCTGATGAGGAAACAAGACGCGCTTTCGGTCTGCA 8343
QY 3195 CTTGGAACCAAGAAAGGGGTGACCATTTAGGAGTATCCCCGGAAGTTTGAACCGTGGT --- 3251
D8 8344 CTTGGAATAGCAAGAGGAGACATCAAGACACCCCGGAAGGACAGAAAGATGCTG 8403
QY 3252 -----CACTAGTTCAGAGCGCTATGCGTCTTTCGATGTCACGTTCCCATGCAAAAC 3305
D8 8404 CTGACACCATGCTGACCGGCTCATGCTTGTCTTGGAAACGTAGAGCTTCCCATGCAATGCC 8463
QY 3306 CACCGGTGCTATTCTACTGACGCAAGAACACACTGAGTCTCGAAGAGAACGTCG 3365
D8 8464 CGCCCACTGCTACACCGCGAACCATTCAGAGCTCTGACATCTCGAAGAGAACGTCGA 8523
QY 3366 ACAATCCAAATTAAGACACGCTGCGAGAAAGTCTTGAAT---GTTCATCAGCGCGC 3422
D8 8524 ACACAGAGGCTTACGACACCTGCTCAACGCTATTTCCGGTGGGATTCGTCGGCAGAA 8583
QY 3423 CCAACGAGAGCATTAACGATGACTTACACTGACCAAGTCCCTACCTGGGGTTCGCCGT 3482
D8 8584 GTAAAGAGAGGCTACGACGACTTTCCTTGACAGCCGCTACTTGGGCACTGCTCTG 8643
QY 3483 ATTGCAAGCACTCAACGCGGTGTTTCAGCCCAATTAATGAGAACGTGGGAGCAAT 3542
D8 8644 ACTTCACCATATGAAACCGTGTTCAGCCGATTAAGATGACAGAGTCTGGAGATGAG 8703
QY 3543 CTGATGATGATTCGATTAGATTCAGAGTCTCGGACAAATTCGGCTACAAATCAGGACCA 3602
D8 8704 CGAGAGACACACATAGAGTACGACTTCCGCCAGTTTGAATAGACCAAGACGAG 8763
QY 3603 CTGCGGATGTCACCAATTCGCTTACATGCTTTTGACCAAGCAGCAGCATGCAAGAG 3662
D8 8764 CAGCAAGCTCAAAATAGTACCGCTACATGCTGCTGAGCAGAGATCATCTCTCAAGAG 8823
QY 3663 ACAATATGAGAAATAGCTATCAGCAATCTGACACCTCGCGCTGCTTTGGCCAAAG 3722
D8 8824 GCACCATGAGATGACATCAAGATCAGACCTGAGACCGGTGAGAGGTTAGTACAAAG 8883
QY 3723 GGTATCTCTGTTGACTCAATGTCCTCCAGGTGACAGTGAACGTCAGTACAGACG 3782
D8 8884 GATATCTTCTCTCGGCAAGTGTCTCCAGGGGACAGGTGAACGTTAGCATTAACAGTA 8943
QY 3783 GAGCATCTGAGATTCATGACCGGTGAGAAAGATCAGAGAGAGTTCGCGTAAAG 3842
D8 8944 GCAACTCAGCAACGTATGCAACAATGGCCCGCAAGATTAACCAAAATTCGTGGAGCGG 9003
QY 3843 AGGATCTTGTTCCTCCACCGCTCATGGAAGCTGTAAATGTCACGTTTACGATCACT 3902
D8 9004 AAAAATATGACCTTACCTCCGTTACGCTGAGAAAGATTCCTTGCAACATGACACGCTC 9063
QY 3903 TGAAGAGACGTCGCGGATACATAACATGACACAGGCGCACACGCGTAAATAGT 3962
D8 9064 TGAAGAAACACCGCGCGCTACATCACTATGCAACGCGCGGACCGATCCCTATACAT 9123
QY 3963 CTTATCTGAGAGAAAGCTCAGCGAAGTGAATTAAACACCTTTGCGAAGAAAGCTCA 4022
D8 9124 CTTATCTGAGAGAAATCAATCAGGAAAGTTTACGGAAGCCACATCCGGGAAGAACTTA 9183
QY 4023 CTTACGAATGTAATGTGGGCACTACAGCACAGTATGTGAGACCGGAAACGAAGTGA 4082
D8 9184 CGTAGAGTGAAGTGGCGGCTATTACAGACCGGAACGTTTACGACCGGTACGGAATCA 9243
QY 4083 ACGGCTGCACTAAAGCAAAACAGTGCATTGCTCAAGAGGACCAAAAGAAAGGCTCT 4142
D8 9244 CGGGCTGACCGGCATCAAGAGTGGCTGCTCTTAAAGGAGCAAAAGAGTGGCT 9303
QY 4143 TCAACTGCGCGGATCTTATTAAGCACACGACCACTCAGTCAAGTAATTTGACATTC 4202
D8 9304 TCAACTGCGCGGATCTGATCAGACACCGCGACACAGCGGCCAAAGGAAATTTGATTTGC 9363

QY 4203 CATTCGCTTGAACACGACAGTGTGCCGGTTCGGTATGCTCACAGCCTTACAGTCACGA 4262
D8 9364 CTTTCAAGCTGATCCGAGATACCTGCAATGCTGCTCTGTTGCCACAGCGCCAACTAGTAC 9423
QY 4263 AGTGGTTCAAAAGGATCACCTCCACCTGACCTGCATGCAACCAATTCGTCGAACGA 4322
D8 9424 ACGGCTTAAACATCAGCTCCAAATTAAGACACAGACATCTGACATTCCTCACCA 9483
QY 4323 GAAATTTGGGCTCGGAGCAGACGCAACAGCAATTTGATTTACAGGCTTACATCCAGA 4382
D8 9484 GGAACATGAGGGGCAAAACCCGGAACCAACTGAATGATTCATGGAACAGGTTGAA 9543
QY 4383 ATTTTCTGTTGGGCGAAGAGGCTGAGTACGTAATGAGGTTAACATGAACCACTGAG 4442
D8 9544 ACTTACCGCTGACCGAGATGCTGTAATATCATATGAGGCAATCAGAACAGTAAGG 9603
QY 4443 TCTGGGCCAGAGTGGGACACAGGCGACCCCACTGATGAGCGCATGAGATCATATCC 4502
D8 9604 TCTATGCCCAAGAGTGCACACAGAGACCTCAGCGATGGCCACAGAAATAGTACAGC 9663
QY 4503 ACTTATTCATCGGCATCCAGTCTACCTGTCATGTCGTGTGTGTGTGCTTCTGCTA 4562
D8 9664 ATTATCATTCATCGCATCTGTGTACCACTTAAAGCGTGCATCAGCTGTGTGCGA 9723
QY 4563 TCTGTGAGCACTGCATCATCAGACAGCTTGACATCGCCAAAGCAAGAAAGACTGCTGA 4622
D8 9724 TGATGATTTGGGCTACCTGTTGACAGATTAATGTGCTTAAAGCGCGCGTGAAGCTGTA 9783
QY 4623 CGCCATACGCGCTTGACACGAGCAACGAGTACCACAGCATTAAGCGTTTGTGCTGA 4682
D8 9784 CGCCATATGCCCTGCGCCCAAAATGCCGTGATTCAACTTGCCTGCGACTTTTGTGCTGTG 9843
QY 4683 TTGGGCCAACCAACGCTGAAACATTTGAGAAACTTTGAACATCTGTGTTTAAACAAC 4742
D8 9844 TTAGTGGGTTATGCTGAAACATTCACGAGACAGAGTTCATTATGTCGAACAGCC 9903
QY 4743 AACGTTTCTGCGGACAGTGTGACATTCCTCTGCGAGCGCTTGTATTCGTTCCGCT 4802
D8 9904 AGCGTTTCTTGTGGTCAAGTGTATCTCTGCGCGCTGTCTGTCTTAATGGCT 9963
QY 4803 GCTTTTCATGCTGATGCTCTTTTATATGTTGTTGAGGCGTCTGCTGGGAAAGGTAGACG 4862
D8 9964 GTTGCTATGCTGCTGCTCTTTTATGTTGTTGCGCGCTACCTGCGCAAGGTAGACG 10023
QY 4863 CTTTGAACATGGAACCACTGTGTGCAAAATGTTCCGGGATTCCTGTATTAAGCGTTGCTG 4922
D8 10024 CTTACGAACATGCGACCACTGTTCCAAATGTGCCACAGATACCGTATTAAGCACTGTG 10083
QY 4923 AACGCGAGGTTACGCGCACTTAACCTGAGATCAAGTGTCTCATCGGAATTAACAC 4982
D8 10084 AAAGGAGAGGTAGCGCCGCTCAATTTGAGATTAATGATTAATGATCTCTCGAGGTTTGC 10143
QY 4983 CTTCACTTAACAGAGTACGTGACCTGCAATTCACACAGTCAATTCCTTACACCAAG 5042
D8 10144 CTTCCACCAACCAAGATTAATTAATGCAATTCACCACTGTGTCCTCCCTCTAAAG 10203
QY 5043 TTAATGCTGCGGTCTCTGAGTGCAGAGCATCTTAAGGCGGATTAACATGCGCG 5102
D8 10204 TCAATGCTGCGGTCTCTGGAATGTACGCCGCTGACGACGATTAATCTGCAAG 10263
QY 5103 TTTTGGCGGTGCTACCTTTTCAATGTTGGGAGGCGACATGCTTCTGACAGTGA 5162
D8 10264 TCTTTGAGGAGGTATACCTCTTCAATGTTGGGAGGAGACCAATGTTTGTGCAAGTGA 10323
QY 5163 ACAACAACTGAGTGAAGCGTGAAGTTCGCTCCAGCTGACATTAATGATACGAG 5222
D8 10324 ACAGCAGATGAGTGAAGCGTGAATGTCAGTAAATTTGCGACATGACACCGCG 10383
QY 5223 TCGCATTAAGTTTACACAGCTGTCTTGAAGTGTGCGCTGCTGTATGATAGGCAACA 5282
D8 10384 AGGCGATTAAGTGTCAATGCTGCGCGATGAAGATGAGATGCTGTATGCTAGCGGAACA 10443

5283 CCACCCGCGACCTGGATACGTTTGCATGCGCTACGCCAGGCTCTCCACGGGACCTGTA 5342
10444 CTACACAGTTTCTTACATGATGTAAGCGAGTCAACACGAGAAAGCTTAAGACCTGA 10503
5343 AGGTATAGCAGGCGGATATAGCGCGCTTTTCCACCTTTTACCATTAAGTTCGTCATTA 5402
10504 AAGTATAGCTGAGCAATTTAGCATTTGTTTACACCATTTGATCAACAAGGTCGTTATCA 10563
5403 GAAAGGGGCTTTTACAACTACGACTTCCCTGAGTATGAGACTTATGAAACAGGAGCGT 5462
10564 ATCCGGGCTGTGTATCACTATGACTTTCCGGAATACGAGACGATGAAACAGAGCGT 10623
5463 TCGCGCATATTCAGCAATCTCGCTTGTATGATGATACAGACATATAGCCGACCTGACATAC 5522
10624 TTGAGACATTCAGCACTACCTCTTGTACTAGCAAAAGACTTATCGCAGACACAGACATTA 10683
5523 GGGCTCTGAGGCTTCTGTCAAGAACATCCAGTCCCTTACACCCAGACAGATATCAGGT 5582
10684 GGGTACTAGGCTTCTCGCAAGAACATGTCATGTCCTGATACAGAGCGGATCTGAT 10743
5583 ATGAATGTGAAAGAACATCAGACGACCCCTGCAAGAAACAGACACTTTGATGTA 5642
10744 TCGAATGTGAAAGAACATCAGACGCGCCCTGCAAGAAACGCGCTTTTGGGTGCA 10803
5643 AAATTGAAGTGAGGCTCTGCGAGCGCTTACTGCTTACCGGACATCTCCATCTGCA 5702
10804 AGATTGAGTCAATCCGCTTCAGCGGTGATGCTGCTCATAGGGAACATTTCCATTTGTA 10863
5703 TTGACATCTCTGATGACGCTTTTGTGATATCATGAAATCACCAACATTTTGAAGTTA 5762
10864 TTGACATCTCTGACGCTTTTATCAGACATCATGATGACACATCTGCTCACACAGTCA 10923
5763 GCTGACAGTACAGACGCTTTTATTTTGTGAGACTTTGTGTGCTTCTTCAACATTAAGT 5822
10924 AATGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10983
5823 ACAAAGCTGACAGGAGGAGGACATTTGCACTTCTCCATCTCAAGACGCTGTTTGA 5882
10984 ATGTATCCGACCGCGAAGGACATGCGCTGATCATTTGATGATGATGATGATGATGATGAT 11043
5883 AGGAAGCGACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5942
11044 AAGATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11103
5943 GCCCAAGCAAAATTTTATGATTTGCTGATGCGGAGAGGATGCACTGCAATGCTGAT 6002
11104 GCCCAAGCGAATCTTATGATGCTGTGTGTGATGATGATGATGATGATGATGATGATGAT 11163
6003 GTAAACCAACCGGACGACATATTTGAGAACCAATTAAGTCAACCAAGATTTCCAG 6062
11164 GCAACCAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11223
6063 CGGCAATTTCCAAACATTTGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6122
11224 CGGCATCTCAAAACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11283
6123 TCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6182
11284 TATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11343
6183 GACTGAGCGGACGACCTGA-CATAGCGGTAATGATGATGATGATGATGATGATGATGATGAT 6241
11344 GACCGTACGCGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11403
6242 GCATATGCGACGCGCT-----TGAACCT 6268
11404 GCATATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11463
6269 AAAATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6318
11464 AAAATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11513

RESULT 12
ACU38305
LOCUS 11663 bp RNA linear VRL 30-JAN-1996
DEFINITION Sindbis-like virus isolate S.A.AR86, complete genome.
ACCESSION U38305
VERSION U38305.1 GI:1125069
KEYWORDS
SOURCE
ORGANISM
Sindbis-like virus
Sindbis-like virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
Alphavirus; WEEV complex.
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ACU38305
Sindbis-like virus isolate S.A.AR86, complete genome.
11663 bp RNA linear VRL 30-JAN-1996
Sindbis virus mutations which coordinately affect glycoprotein
processing, penetration, and virulence in mice
J. Virol. 63 (4), 1619-1623 (1989)
89178848
2926866
2 (bases 1 to 11663)
Simpson,D.A., Lin,S.C. and Johnston,R.E.
Submitted (11-OCT-1995) Dennis A. Simpson, Microbiology &
Immunology, University of North Carolina, Rm834 FLOB, CB7290,
Chapel Hill, NC 27599-7290, USA
Location/Qualifiers
1..11663
/organism="Sindbis-like virus"
/mol_type="genomic RNA"
/isolate="S.A.AR86"
/db_xref="taxon:4419"
/note="from South Africa."
1..59
60..7562
/codon_start=1
/product="nonstructural protein"
/protein_id="AA86135.1"
/db_xref="GI:1125070"
translation="MEKVVNVADVDPGSPFVVOLOKSPFPEVVAQVTPNDHANARA
FSLASKLIELEVPPTTATLIDGAPARMSELOHVCMBRSDPDMKVASKL
AEKAKITNKILHEKIDRLVLDTPDAETSLCEHNDVCTNTRAEVSWDDVYNAP
GTHHQAKGVRTLYWIGFTTQFMFSMASSYPAYVNTNMDERKLEKRLGLSTKL
SEGRGKLSIMRKELKPSRVSYSVSTGLPEHRAISQSWHLPSVFLKQOSYTCR
CDTVSCGCVVKKITLISPGITGEVKNVNSGFLKCTVDTVKEGRVSPVCTY
IPATICDQMTGIMATDISPDQOKLVGLNRIYVNGKTNNTNMYLPIAOGF
SKMKERKEDI.DNKMIGTTERKLTTCGLMAFRKKXHSFYRPGTORTIVRPSFA
FPMSSVWTTSLPMSLRQMKALDPKSEKILQVPEELMEAKAFEDQESRBEKL
REALPPLVADGIEBAAEVCEVEGLQDVAALVETPRGRVRIIPQANDMIGQYIV
VSPISVLNNAKLAPAHPLADQVKIITHSGRGRVAVEYDAKVLMPAGSAVPEFLA
LSEATLVYNEREFVNRKLYHIAHGPARKNTEEOYKTKALEVETVFPVDDKRCV
KKEASGLVLSGELTNPYHELALGLKTRPAVYKVETIGVIGPGSGKSAIISKTV
TABRLNLSGKENEKEIEADVLRIKMOITRSKYDVSVMNCHAVELVYDEARCH
AGALLALAIVRPKKRVLLCGDPKQCGFFNMOLKVHNPHEKDICTFTFKFSRRC
TOPVTAIVSTLHYGKKNTNPCKNIEIDITGATKPKPGDIILTFEGMWKQLOIDY
PGHEVMAAASQGLTRKGVYAVRQKVENPLVATSEKVNVLRTBEDRLVTKLOGD
PIKQLTNPVPGNFQATIEDWEAHKGI.IAALNSPAPNTNPFSCIKTNCAKALAEPLI
ATAGILVILGCOMSEL.PQPADKXHSALVLDVLCIKFEGMDLTSGLFSKOSIPLTV
PADBARPRAHNDNSPGRKXYGDYAAVAELSRPVPOLAKGTOLDIQTERTVISA
QHNLVPPVNRNLPALVPEHKKQGPVEKLSQRHNSVLVISEKTLAPKRIEWA
PIGIAGADKNVNLAFEPPOARVLDVLTNGTKYRNHFOQCEHAAITLTSRSLN
CLNPGGTLVNSGYADNSEDVVTALARKVRSAAPECVSSNTEMYLTFROLDNS
RTROTFPHNLNVLSVYEGTRDVGAPSRTRKRENTADQOEVAANAAPLQGE
GVCAIYKRMNKSPTDSATENGTAULTVCOGKQVYHANGPRGRHPEALIKLONAY
HAYDVAIYHNRKSVAILPILSTGIYAAGKDELVSINCLTTLALDRTPADVITLCKDK
WKERIDAVLOKESVTELDKDEMDIEBLVITHPDSCKKRGKSTTKGKIYSFEGT
KEHOAAOMAEIKVLPVDOESNEQLVILIGETMEALREKCPVDHNSAPPPTLP
LCWMTAMPERRVHRIRSNNVKEVTCSSPPLKIKYIKVQKQCTKVLFNHPTPELP
ARKYIEAPEQVPAQAEAPVAVPTPAAANTSLADVDILDMDESSEGLSS
PESGDNRQVAVDAVAVOEPVAVPRLKXMAALAAAROEPTPASTSADSL
HLSDGVLSIFSGSIDBEMARLAAOPASTCPTDVMASRPSFGDIEIESTRTES
EPVLPSTREBEVNSIISRSVAFPPRKORRRRSRTREYCLGVGVIISTRTGRC
HLQKSVLQNOVLTPTLERNVLETVIAPVLDLTSKEQDKLQVOMPTPEANSRTQSRK
YHENOAITTERLLSLGLRLVNSATQPCYKITYPKPSVSSVPANYSDPKPAVAVCNV
YIHENYPTVASVOYITDEYDAYLDMVDGTVALCTATPCPAKLRSYPRKHEVRAINRS

	mat_peptide	product="nspl"	60.1679	AVSAMSNTLQNVLIATKRCNVTQKRELPTLDSAPFNVECFKPKYACNDYEMEEPRK KPIITTEFVTAIVARLKGPPAAALFKATHNRLVLQSPMDRPMWMDKDVKTVPGRK HTEPRPVQVIOQAEPPLATAVLCGIHRELVAHRLVAVLPINHTLPFMSADPFAIAIE HFHKGSDVLEETDIASPDKSPQSDAMALVTGLMIEDLVGDQPLDILIECAFGBISSTHLP TGTEFKFGAMKMSGMFLPLFNATLVNIAASRVLEELRSTKSCAATFGDNIHGVVS DKMEARCATWLMNEVKIIDLAVIGERPPYCGGTLIDSVTSTGRADPLIKRLFKLG KPLADQEDQEDRRRALLDETAKMFRGVIDTLVAVATREVDNITPVLLATPFAQ SKRAFOAIRGEIKHLYGGPK"
	mat_peptide	product="nspl"	1680.4100	
	mat_peptide	product="nspl"	4098.5729	
	mat_peptide	product="nspl"	5730.7559	
	mat_peptide	product="nspl"	7608.11345	
CDS		codon_start=1		
		product="structural polypeptide"		
		protein_id="AA08136.1"		
		db_xref="gi:1125071"		
		translation="MNRGSEFMMLGRBPPEPTAMRPRRRROAAPMEPARGLASQIQO LTVTAASLVIGQATRPQTPRRPPEPRQKQAPKPPKPKTKTEKKKQPAKPKPK RORALKLEADRLFDVKNEDGDVIGHALMFGKWKPLHYVGTIDHPLVSLKTKTKSS AYDNEPQALPVNMRSEAFPTYSSEHPEGRVNHGAVOYSGRFTIPRGVGRDQSGSP IMDSGRVAVLQGADEGRTTALSVYTANSSKGTITLTTPTEGSEKAPLVYMCLL GNVSPCNRPPTCYTREBSRALDLLEENVNEADYTLNATLIRCGSSGRSVYTDLP RLTSLVPLGTCYSCCHTEPCSPKIKIEQVMEBADNTIRIOTSAPGYDQSGAASNKY RYSMLIEDHTVKEGTMDDIKISTGSPGRSLYKGYFLAKPPDSVTVSIASSNSXK SCIMARKIKPKFVGRREKYLDPVHGKIKPCVYRLKETTPAGVITMRPGHATYSLS EESGKYVAKPKPGKNITVECKGDYKGTITPTLITGCAIKOCAYVSDOKWVF NSPDSIRHADHTAGKHLPRFKLIPSTCMVVAAPRVNVEFKISLQTDHDLTLT TRIRGANPEPTTEMIIGNTAVNFTVDKGLIEIGNHEPVRVAQESAPGDPHGMPE IVQYVHRHPVYTIILAAVAANMIGVYVALACCAKARECLTPVALAPAVIPTEL ALILCVSANALETETEMSYLYMSNSOFPVQOLCIPLAAVVLRRCCSCPLFVAVG AYLAKVAVEHATVTPVNPQIPYKALVERAGYALNIEITMSEVLPSTNOEYITCKP FTTVAPSPKRVCGSLGECOPAAHADYTCXVGIVYPRMWSAQOCFDSSENSOEAAVV ELSDCATDHAQAIKHTAANKVGLRIYVGNSTSPFLVYVNGVTPGTSKDLKVIAGSI SALTTPDHKVAIRKGLVNTYNDPEYGMKRGAGGDIQATSLISKDLASTDIKLLPK SAAKHVAVPYTQASGFEWKKNSGRPLQETAPEFCKLAVNPLRAVDOSYGNIPISIDI PNAPFIRITSDAPLIVSTVCIDVSECTYSDFEGMATLQVSDREGCQVSHSSTATLO ESTVHLEKGAIVTHFSTASPOANFIVSLCKKTTCAECKPRADHIVSTPHKNDQGEF QALSKTSMWSLFPALFGASSLIIIGLMIFACSMWLTSTR"		
	mat_peptide	7608.8399		
	mat_peptide	product="capsid"	8400.8591	
	mat_peptide	product="E3"	8592.9860	
	mat_peptide	product="E2"	9861.10025	
	mat_peptide	product="E6"	10026.11342	
	mat_peptide	product="E1"	11346.11663	
BASE COUNT	3'UTR	11346.11663		
ORIGIN	3291 a	3105 c	2907 g	2360 t
	Query Match	36.9%	Score 2396	DB 14
	Best Local Similarity	64.6%	Pred. No. 0	Length 11663
	Matches 3751	Conservative 0	Mismatches 1985	Indels 74
			Gaps 9	
QY	580	ACTGACGGTATGAAGCGGAGCGATATATTTCTCATCGGAACAGGCCAAGCTACCTTC	639	
DB	5707	ACTGCTTAACCGGGGTAGTGGGTAACAATATTTTCAGACGGAACAGGCCCTTGGGACCTTGC	5766	
QY	640	AACAGAAATCAGTACGTCATGTAAACTACAAAGAACTATTTGGATCGGCGCCTTCATG	699	
DB	5767	AAAAGAAATCCGTTCTGCAAGAACACGCTTACAGAACGACCTTGAAGCGCAATGTTCGG	5826	
QY	700	AGAAGTATTACGCCCGCGCGCTCGATCTCGAAAGAGAGAANAATGTTACAGAAAGAACTCG	759	
DB	5827	AAAGAAATTCAGCCCCCGGTGCTGCGACAGTGCAGAAAGAGAACAGCTCAAACTCAGGTACC	5886	

QY	760	AATTATGGCCTTCGAAAGAAATAGAAAGCAAGTATCAATCAAGAAAGTAAAGAAATATGA	819
Db	5887	AGATGATGCCCAACGAAAGCCAAACAAAGCAGGTACAGTCTCGAAAGATGAAAAACAGA	5946
QY	820	AAGCAATTACAGCGGAGCACTCAATTTCTGGATTGGGACATATCTATCATCAGAAAGTGA	879
Db	5947	AAGCATTAACCACTGAGGCACTGCTTTCAGGGCTACGACTGTATTAACCTTGCACAG---	6003
QY	880	ATCTGTGAGTGTTCAGAGTCAATTATTCCTGTACCAATCTACTCGTCAACGGTAATTA	939
Db	6004	ATCAGCCAGAAATGCTATTAAGATCACTACCCGAAACCATCGTATTTCCAGCAGTGTACCG	6063
QY	940	ACAGTTTACATCTGCAGAGGTCCGGTTAAACGTGCAACTTACGTTATCCAAAGAAATT	999
Db	6064	CGAATCTACTGCAGCCCAAGTTTGTCTGAGCTGTTTGTAAACATCTCTGCATGAGAAAT	6123
QY	1000	ACCCACAGTAGCCCGATTAATTGTATTAACAGATGAAATACAGTGCATCTTGACATGCTGG	1055
Db	6124	ACCCGACGGTAGCATCTTATCAGATCACCGACGAGTACGATGCTTACTTGGATATGTGAG	6185
QY	1060	ACGGGCATCGTGTCTGTAGATACAGCACATTTTGTCTCGGCTAAACCTGAAAGTACCC	1119
Db	6184	ACGGGACAGTGCCTTGCTAGTACTGTCAACTTTTGTCCCGCCAGCTTGAAGTTATCC	6243
QY	1120	CAAAGAACATATGCTATTTGCAAGCCAGAGATTAAGTCAAGCCGCTCCCATCCGCTTATACGA	1179
Db	6244	CGAAAGACAGAGATATAGAGCCCAACATCCGAGTGCAGTTTCCATCAACGATGACGA	6303
QY	1180	ATACATTCAAAAATGATTTGGCTGAGCTACTAAAGGAATTGCAACGTTACCCAAATGC	1239
Db	6304	ACACCTTCCAAAACGTGCTCATTTGCCGGACTTAAAGAACTGCACCTCACACAAATGC	6365
QY	1240	GAGAAATTACTGTCTTATGATTCCGGCGCAATTAAATGTTGATTGTTCAAGAAATACGAT	1299
Db	6364	GTTGAACTGCCAAACACTGGACTCAGGACATTTCAACGTTGATGCTTTGAAAAATATGAT	6423
QY	1300	GCAATGATGAGTACTGGGATACCTTTGGCGATAACCTTATCCGCTAATCTACAGAAAC	1355
Db	6424	GCAATGAGCAGATATTTGGAGAGATTTGCCCAAGCCCAATTAGATCATCTGAGTTTCC	6483
QY	1360	TTAGCCATATGTGACAAAGCTGAAAGGGCCGAAAGCAGCAGCATTTGTTCCGAATACTC	1419
Db	6484	TTACCGCATAGTGGCCAGACTGAAAGGCCCTTAAGCCCGCAGCTGTTCCGAAAGACGC	6543
QY	1420	ATAATCTAAAAACGTTGCGAGAGATATCCAAATGATCAATTTGTCATGATCTAAAGAG	1479
Db	6544	ATTAATTTGGTCCCATGCAABAAGTCCCTTAGATAGATTTGTCATGACATGAAAAAG	6603
QY	1480	ATGTCAAAAGTATCTCCCGCACAGAAACATACAGAGAGCGGCTTAAGTGCAGTTATTC	1539
Db	6604	ACGTAAAGTTTACACTGTGGCACAGAAACACACAGAAAGAAAGACCGAAAGTAAAGTAAC	6663
QY	1540	AGCTGCGAGATCCCTTGGCTACCGCTTACCTTTGGGGATTCATCTCCGGGAATTAGTCGTA	1599
Db	6664	AAGCCGAGAAACCCCTGGCGACCGTTTACCTATGGGGATTCACCGGGAGTTAAGTCCCA	6723
QY	1600	GACTGAATGGCGTGTCTTGTGCGAAATATCACTACTCTTTCGACATGTCAAGCGAAGATT	1655
Db	6724	GCGTTACAGCCGTTTGTGCTACCCAAATTCACAGCTCTTTGAGATGTGCGCGAGACT	6785
QY	1660	TTGATGCAATTAATGCTGTAACATTTCCACACGGCGACCCAGTATGGAAACGACATCG	1719
Db	6784	TTGATGCAATCAATGACAGAACACTTCAAGCAAGGTGACCCGCTACTGAGACGGATATCG	6843
QY	1720	CGTGTGTTGATTAAGGCGAAGACGATATCCGCAATTCGCGCGTTGATGATCTCTGAGG	1779
Db	6844	CCTCGTTTGACAAAGCCAAACGACGCTAATGGCGTTTACCGGCGCTGATGATCTTGGAG	6903
QY	1780	ACTTAGTGTGCAACCAACGCTCTTATGATTGATAGAGCGCGGCTTCGGCATATATCAT	1839
Db	6904	ACCTGGGTGTGACCAACCACTACTCGACTTGATGAGTGGGCGCTTTGGAAATATATCAT	6965
QY	1840	CTGTGCACTTACTACAGAAACGAGGTTAAATTTGTGTCATGATGAATCCCGTATGT	1899

Db 6964 CCACCATCTGCCCCAGGGTATCCCGTTCAAAATCGGGGCGATGATGAATCCGAATGT 7023
Qy 1900 TCTTAAGCTGTTTGTCAACACATAGCAATATCATATGCTGACGAGACTACGTG 1959
Db 7024 TCTCAAGCTCTTGTGTCAACACATGTTGAAATGCTATGCGCAGAGATTTGAGG 7083
Qy 1960 AACGGTTAACACATGAGCGTGCAGGCGCTCTATCGCGACGATTAACATAGTGAATG 2019
Db 7084 AGCGCTTAAACGTCGAATGTGACATTTATGCGGACGACAACTATATACGAG 7143
Qy 2020 TCGTCTCGACACCTTGATGCGGAGAGATGCGCACTTGCGTGAACATGGAATGAAA 2079
Db 7144 TAGTATCTGACAAAGAAATGCTGAGAGGTGTCACCTGCTCAACATGAGGTTAAGA 7203
Qy 2080 TTATGATGACGATTATGTTATCAAGACCTACTTCTGTGGGGATTTATCTGCTGG 2139
Db 7204 TCATGAGCGCATGCTCGCGAGAGACCACTTACTTGTGCGGTGATTCATCTTGCAAG 7263
Qy 2140 ACCAGATTAACAGGACAGCTGTCAGAGTGCAGACCTCTTAAGAGCTTTTAAAGTTG 2199
Db 7264 ATTCGTTTACCTCCACAGCGTGTCCGCTGGGAGCCCTTGAAAAGCTGTTTAAAGTTGG 7323
Qy 2200 GAAAAACAATGCGCATGATGATACCAAGCTGCGACCGCGCGCATGATGATG 2259
Db 7324 GTAACCCGCTCCACCGACGATGAGCAAGACGAAGACGAAGACGCGCTGCTAGATG 7383
Qy 2260 AAGCAATCGATGGAACAGAAATTTGAATTTACGAGAGATTAAGTGAAGCCGTGAATCA 2319
Db 7384 AAAAAAGCGGTGTTTGAAGATTAAGACCACTTACAGTGGCCGCGGCACTC 7443
Qy 2320 GATACGATCATATCTGCGAGGCTGATCATCAGCTCTCTGTCACAGTTAGCCGAAGCG 2379
Db 7444 GGTATGAGGTAGCAACAATCACACTGTCTGTGCTGATGAACTTTTGGCCGAGACA 7503
Qy 2380 TTAAAGACTTCAGAGCATTAAGAGGAGCCCAATCACTCTTACGCTGACCTTAATAG 2439
Db 7504 AAAAGCATTTCAAGCCATCAGAGGGGAAATTAACATCTCTAAGGTGTCTTAATAGT 7563
Qy 2440 TGACCTATTAGA-----CAGCACCTAACCCAGCGGCAAGATGTTTCCA 2482
Db 7564 CAGCATATGATCAATTTCTGATCTAATAACACACACACCACTGATATGAGATTTCT 7623
Qy 2483 TACCTCAGCTGAATTTTCCACAGTTTACCTTAACAAATCCGATGAGTTTACGAGATCCA 2542
Db 7624 TTAACATGCTGGCGCGCGCCCTTCCAGCCCTCCACTGCTGATGAGGCGCGGAGAA 7683
Qy 2543 AACCTCTAGGCGCGCTGAGAGCGGTTTGGCCCGCTGGCTGCTCAATTCGAGAT 2602
Db 7684 GAGGCGAGGCGCGCCGATGCTGCGCCGCAATGGGCTGGCTTCCCAATTCAGCAACTGA 7743
Qy 2603 CTTAGAGAGTGAATGATCACTTTGACCTTTCAAAACAAC--GATCACTTAATCCGCGCAG 2660
Db 7744 CCACAGCGGTAGTCCCTAGTACTTGAACAGGCAACTGACCTTAAACCCCAACCCAC 7803
Qy 2661 GTCCACCGCCAAAGAAAGAAAGAGTGTCTTAAGCCAAACCTTACGAGCTTAAAGAA 2720
Db 7804 GCCCGCGCGCGCCAGAAAGAGAGGCGCCAAACCAACACGAGCCGAAAGAAACCA 7863
Qy 2721 AGAAGCAGCAAGCCAAAGAGAGAAAC--GCAAGCTTAACCAAGGAAACGACACGTA 2777
Db 7864 AAAACAGAGAAAGAAAGAAAGAAACCTGCAAAACCCGAAAGAGACGCTTA 7923
Qy 2778 TGTGTATGAAGTTGAGAGTGCAGACATTTCCGATC--ATGCTGAACGCGCAATGA 2834
Db 7924 TGGCATTAAGTTGAGGCGCGACAGTCTTGAAGTCAAAATAGAGACGAGATGTCA 7983
Qy 2835 ATGATATGCGCTGCTGCGAGAGAGGCTGATGAACCACTCCAGTTGAAGAGAAA 2894
Db 7984 TCGGCGACGCACTGCGCATGAGAAAGGTATTAACCACTCCACGTTGAAGGAACTTA 8043
Qy 2895 TTGATTAATGAGCAATTAAGCGCGCGTGAATTAAGAAAGGCTAGCATGTACGACTTGAGT 2954
|||||

Db 8044 TTGACCACTGTGCTATCAAAAGCTCAAAATTCACCAAGTCTGACATACGACATGAGT 8103
Qy 2955 ACGGAGAGTTCCTCCAGAAACATGAAATATGACACGCTGCACTATACACGAGCAAAACAC 3014
Db 8104 TCGACAGTTTCCGGTCAACATGAAAGTGAAGGCTTACCTACACGATGAACCCCTG 8163
Qy 3015 CGGCTTCTTAACATGAGCACCAGCGCGAGTCCAGTATGAATGAGAGATTTTACCGTAC 3074
Db 8164 AAGGTTCTAACAATGAGACCAACGAGCGGTGCAATATGAGAGGACGATTTTACATCC 8223
Qy 3075 CGAAGAGATGCGCGGAGAAAGCGACAGCGGAGACCGATCTTGACAAACAGAGCGAG 3134
Db 8224 CCGCGGAGTGAAGAGCGAGAGAGACAGTGTCTCCATATGATTAATCAAGCCGCG 8283
Qy 3135 TTGCTGATATGTTCTAGAGAGTGCAAATGAGGACACCGCTTACAGTGTCA 3194
Db 8284 TTGTGCGATAGTCTCGAGAGGCTGATGAGGAAACGAACCGCTTTGCGTGTCA 8343
Qy 3195 CTTGAACCAAGAGGAGTACATTAAGGATACCCCGAAGTTCGAAACCGTGT-- 3251
Db 8344 CTGCAACCAAGAGGAGAGCAATCAAGACACCCGGAAGGACAGAAAGTGTCTG 8403
Qy 3252 -----CACTAGTACAGCGGTATGCTGCTTTCGAATGTCAAGTTCCATGCAAAAC 3305
Db 8404 CTGCAACCACTGTACCGGCTATGCTTGTGGAAGAGTGAAGTTCCATGCAATGCC 8463
Qy 3306 CACCGTGTGTATTTCACTGACCGCAAGACCAACTGACAGTGTCTGAAGAAAGTGTG 3365
Db 8464 CGCCCACTGTACACCGCGCAACCAATCCAGAGCTTCGACATCTCGAAGAAAGCTGA 8523
Qy 3366 ACAATCCAAATTAAGACACGCTGCTGAGAAAGCTTGAAT--GTCATCACGCGCGC 3422
Db 8524 ACACAGAGCCTTAGACACCTGCTCAACGCAATTTGCGGTGCGATGCTGCGGAGAA 8583
Qy 3423 CCAAGAGAGATTAACGATGATCACTGACACGCTTCCATCTGCGGATTTGCGCGCT 3482
Db 8584 GTAAAGAGAGCTACGACGACTTTACTTGAACAGCCGCTACTTGGGACAGATGCTGT 8643
Qy 3483 ATTGCAAGACTCAACGCGCTGTTTACGCCCAATAAAAATGAGAAAGTGTGCGAGAA 3542
Db 8644 ACTGTACCAATAGAAACCGTGTGTTAGCCGATTAAGATGACAGAGTCTGGAGATGAG 8703
Qy 3543 CTGATGATGATGATTAAGATTAACGAGCTGCGGACCAATTCGCGCTAACAATCAGGAGCA 3602
Db 8704 CGAGACACACATAGACATACAGACTTCCGCCAATTGATGATGACCAAGCGGAG 8763
Qy 3603 CTGCGGATGTCACCAAAATTCGTTACATGCTTTGACACAGACATGACATCAAGAG 3662
Db 8764 CAGCAAGCTCAAAATTAAGTACCGTATATGCTGTCGAGCAGAGATCATCTGTCAAGAG 8823
Qy 3663 ACATATGAGAAATATAGCTATGACACATTTGACCTGCGCTGCTTGGCCCAAAAG 3722
Db 8824 GCACCATGAGATGACATCAGATCAGACCTCAGAGACCGTGTAGAGGCTTGTACTAAG 8883
Qy 3723 GGTACTCTCTGTGCTCAATGCTCTCAGGTGACAGTGAACGCTGATCAAGCGC 3782
Db 8884 GATACCTTTCTCTGCGAAGTGTCTTCAAGGAGACAGCTTAACGTTAGCTTACGAGTA 8943
Qy 3783 GAGCATGTGAATTAATGACACCGTGAAGAAAGATCAGAGAGAGTTTGTGCTAGAG 3842
Db 8944 GCAACTCAGCAAGCTATGACATAGGCGCGCAAGATTAACAAATTCGTTGGAGCGG 9003
Qy 3843 AGGAGTACTGTTTCCACCGCTCCATGAGAAAGCTGTGAAGTGCAGGTTTACATCACT 3902
Db 9004 AAAAATATGACCTTACCTCCGTTACGCTTAAGAAATTCCTTTCAGAGTATACACCGTC 9063
Qy 3903 TGAAGAGAGCTGCGCGGATACATAACATGACAGGCGCGGACACGCGATTAAGT 3962
Db 9064 TGAAGAAACACCGCGGCTATCACTATGACAGGCGCGGACCGGATCTTATACAT 9123
Qy 3963 CTTATCTGAGAGAGCTGACAGCGAGTGTATTAACCACTTCTGCGAAGAGCTCA 4022
Db 9124 CCTATCTGAGAGATTCATCAGGAGAAAGTTTACGGAAGCCACATCCGGGAAAGAACTTA 9183
|||||

QY	4023	CTTACGGAATGTAATGTGGGACCTCAGCA	CAGTATTCGTGAGCAGCGCAAGGAATGA	4082
Db	9184	CGTAGCGTGAAGTGGCGGATTTACAGAC	CCGGAACCGTTACGACCCGTGACCGAATTA	9243
QY	4083	ACGGCTGCACCTAAAGCAAAACAGTGCA	TTCTCCTACAGAGCCGACCAAGAAATGGTCT	4142
Db	9244	CGGGCTGCACCGGCATACAGACAGTGGGT	CGCTTAAGAGCCACCAAGAAAGTGGTCT	9303
QY	4143	TCAACTGCGCGGATCTTATTTAGGCAC	ACACATCTCAGTGCAGAGTAATTGCACTTC	4202
Db	9304	TCAACTGCGCGGATCTGATCAGACACGC	CCACACACACGCGCCAAAGGAAATTTGCA	9363
QY	4203	CATTCGCGCTTGACACCGACAGTGTGCG	CCGGTTCGTTAGCTCACAGCCCTACAGTACA	4262
Db	9364	CTTTCAAGCTGATCCGAGTAAGTCACTGA	TGATGCTTGTCCACGCGCCGAAAGTATGAC	9423
QY	4263	AGTGTTCAAAGGCATCACCTTCACCTGA	CTGCATTCGACCAACATTTGTCACACGA	4322
Db	9424	ACGGCTTTAAACACATCAGCTTCCATTGA	CACACACCATCTGACATTTGCTCACACCA	9483
QY	4323	GAAATTTGGGCGTGCAGCAGACGCAAC	ACACAAATAGATTTACAGGTTACATCCAGA	4382
Db	9484	GGACACTAGGGGCAAAACCCGGAACCAAC	CACTAATAGATCATCGGAAACACGTTTAGAA	9543
QY	4383	ATTTTCTGTGGGCGAGAAAGGCGTGAGTA	CGTATTCGGGTTAACATGAAACAGTCAGAG	4442
Db	9544	ACTTCACCGTGCACCGAGATGGCGTGAAT	ATACATTTAGGGCAATTCAGAACCGTAAAGG	9603
QY	4443	TCGTGGGCGCCAGAGATCGGCACACAGG	CAACCACTGATGCGCGCATGATCATATCC	4502
Db	9604	TCCTATGCCCAAGAAATCTGCACACAGGAA	ACCTTCACGATGGCCACACGAATTTGTAACAG	9663
QY	4503	ACTATTATCATTCGGCATTCAGTCTACAC	TCTCATCTGTGCTGTGTGCTGCTCTTGCTA	4562
Db	9664	ATTACTATCATTCGGCATCTCTGTGACAC	ATCTTAAGCCGTGGATCAGCTGCTGTGGCA	9723
QY	4563	TCTGTGAGGCATGTCAATCAGACCTTGCA	TTCGCAAGCAAGAGACTGCTGTA	4622
Db	9724	TGATGATTTGGCGTAACTGTTGACAGCAT	TATGTCTGTAAAGCGCCCGTGAAGTGCCTGA	9783
QY	4623	CGGCATACGGGCTTGACCGCAAGCAAGG	TACCCACAGATTAAGGGTTTTGTGCTGCA	4682
Db	9784	CGGCATACGGCTTGAGCCCAATAGCGGTAT	TCCAATTTGCTGAGCACTTTTGTGCTGTG	9843
QY	4683	TTTCGGCAACCAACGCTGTAACAATTTG	AGAAACCTTGTGATTTTAACACC	4742
Db	9844	TTAGTTCGGGCTAATGCTGAAACATTTCA	CCGAGACATGATTAATTATGTCGACACGCC	9903
QY	4743	AAACGTTTCTCTGGGCAAGTTGTGCATTC	CTCTGCGAGCGCTTGTATTTCTCTCCGCT	4802
Db	9904	AGCGGTTCTCTGGGCTCAGCTGTGTATAC	CTCTGCGCGCTGTGCTGTCTATATACGCT	9963
QY	4803	GCTTTTATGCTGATCCTTTTTTATTTG	TGTGCGAGGCGTCTGCTGGGGAAGTGTAGC	4862
Db	9964	GTTCTCATGCTGCTGCTCTTTTATGTG	TTGCTGGCGGCTTACGTGCAAGGTATGAGG	10023
QY	4863	CTTTTGAAATGCGACCACTGTGTCAAA	TGTTCCGGGAATCCCGTATTAAGCGCTTGGTCG	4922
Db	10024	CTTACGGAACATGCGACCACTGTTTCCA	ATGTGCCACAGATACCGTATTAAGGCACTTGTGG	10083
QY	4923	AAAGCGGAGGTTAAGCGGCACTTAACTG	AGATACACGCGTCTCATGCGAATTAAACAC	4982
Db	10084	AAAGGCGAGGTTAAGCGGCACTTAACTG	AGATACACGCGTCTCATGCGAAGTTTTCG	10143
QY	4983	CTTCAACTTAACAGAGTACGTGACCTG	CAAAATTTCCACACAGTCAATCTCTTCCACCAAG	5042
Db	10144	CTTCCACCAACCAAGAGTACATTACCTG	CAAAATTTCCACACATGTGTCCTCTCCCTTAAG	10203
QY	5043	TTAAATGCTGGGGTCCCTCGAGTGCAG	AGGCATCTTCAAAAGCGGATTTACACATG	5102
Db	10204	TCAATGCTGGGCTCTCTTGGAAATGTAC	CGCGCGCTGCACGACATTAACCTGCAAG	10263

OY	5103	TTTTTGCGCGGTGCTACCCCTTTTCATGTGGGGAGAGCCGACAATGCTTCTGTGACAGTAGA	5162
Db	10264	TCCTTTGAGAGGGGTGTACCCCTTCATGTGGGAGAGACAAATGTTTTTGACACGTAGA	1032
OY	5163	ACACAACTAGAGTGAAGGCGTACGTCAGATTCCGTCAGACTGCACATTAATGATCACGCAG	5222
Db	10324	ACAGCCAGATGAGTGAAGGCGTACGTCGAATTTGTCAATGATTTGGCGCATCTGACCAAGCGC	1038
OY	5223	TCGCACTAAAAGTTCACACAGCTGTCTGAAAGTCGGCGTTCGTATATGATTAACGGCAACA	5282
Db	10384	AGCGGATTTAAGTGTCACTACTGCGCGCATGAAGAGGACTGCGATATGTTGTAAGGGAACA	1044
OY	5283	CCACCCGCGCACCTTGATATCGTTTGTCAATGGCGTCACGCCAGGTTCTTCAACGGGACCTGA	5342
Db	10444	CTTACCACTTTCTTAAGATGTATAGCTGAACGAGCTCACACAGGAACGCTCTAAAGAACCTGA	1050
OY	5343	AGGCTATAGACAGGGCGCATATACAGCGGCTTTTTCACCCCTTTGACATTAAGTGGTCATCA	5402
Db	10504	AAGTCATATAGCTGGAACCAATTTTCAGACTTTGTTTACCACTTTGATCACAGAGTGGTTATCA	1056
OY	5403	GAAGAAGGCGTTGTTTCACTACAGACTTCCCTGATATGAGGCTATGAAAACAAGAGCGT	5462
Db	10564	ATCGCGGCGTGTGTGTACACTATGACTTTCCGGATATACGAGGAGATGAACCAAGAGCGT	1062
OY	5463	TCGGCGATTTTCAAGCACTCTCGCTTATGCTACAGACATATGATAGCCCGCATGTACATAC	5522
Db	10624	TTGGAGACATTTCAAGTACCTCTTGACTGTGCAAAAGACTTCATCTGCCAGACAGACATTA	1068
OY	5523	GGCGTGAAGCCCTTCTGTCAAGAACATTCACAGTCCCTTACACCCAAAGCATACAGGT	5582
Db	10684	GGCTACTCAAGCCTTCGCGCAAGAACGTGATGCTCCGTACACGACAGCGCGCATTTGGAT	1074
OY	5583	ATGAATGTGGAAGAACATCTCAGAGACAGCCCTGCAAGAAACAGCACTTTGGATGTA	5642
Db	10744	TCGAGATGTGGAATAAACATCTCAGCGCGCCCATCTGACAGAAACCGCCCTTTGGGTGCA	1080
OY	5643	AAATTTGAATGTGAGCCCTCTGTGGAGCGCTTACTGTGCTTACGGGCATATCCCTATCTGCA	5702
Db	10804	AGATTTGACATCAATCCGCTTCTGAGCGGTGAGCTGCTATACGGGAACATTTCCATTTCTTA	1086
OY	5703	TTGACATCCCTGATAGCCTTTTGTGATCATCTAGAAATACCAACAATTTTGAAGTTA	5762
Db	10864	TTGACATCCCGAAGCGCTGCTTTATTCAGGACATCTAGATGACCACTGGCTCTCAACACTCA	1092
OY	5763	GCTCACAAGTACAGCAGCTGATTTATTTCTCAGACTTTGGTGGTTCTCTAAACATTAAGT	5822
Db	10924	AATGTGATGTCACTGATGTGACTTTATTCACGCACTTCGAGAGGATGGCTTACCTCGAGT	1098
OY	5823	ACAAAGCTGACAGGAGGAGCAATTTGTCAGTTCACTCCACTCCACGACAGCTGTTTTGA	5882
Db	10984	ATGTATCCGACCGGGAAGGACAAATGCCCTGTATCTTGCACTTCAGACACAGCAACCTCC	1104
OY	5883	AGGAAGGACCAACATGTGACTGCCGTAGGACGATTAACATCACTTTTATAGCACAATGCA	5942
Db	11044	AAGAGTGCACATTTCAATGTCTCTGGAAGAAAGAGCGGTGAACATCACTTACGACACCGCA	1110
OY	5943	GCCCAAGCAAAATTTTATATGTTTCGCTATGCGGCAAGAATCCACTGCAATGTCTAAT	6002
Db	11104	GCCCAAGGCGAATTTCAATTTATTCGTGTGTGTGAAGAAAGACAACTGCATATGCAAT	1116
OY	6003	GTAACCAACCGGCGCACCATATTTGGAGAACCACTAAAGTGCACCAAGATTTCCAG	6062
Db	11164	GCAAAACCAACGATCATATTCGTGAGCAACCCGCACAAAAATGACCAAGAAATTCAG	1122
OY	6063	CGGCAATTTCCAAAACATCTTGAACCTGGCTGCTGTGACATGTTTGGGGAGCATCATCC	6122
Db	11224	CCGCACTCTCAAAAACCTTCATGTGAGTGGCGCTGTTTGGCTTTTCCGGGCGCGCTCGTGC	1128
OY	6123	TCATTTGTTGTAAGCACTATAGTGTGTGTCTGAGCTTATAGCTTATAAACACAGTAT	6182
Db	11284	TATTAATTTATAGCACTTATGATTTTCTTGCAGCATGATGTGTGACTTGAACACAGAAAT	1134
OY	6183	GACTGACGCGGCACTGA-CATGACGGTAAATCCATGTACTTCCGAGAGCGTGT	6242

[illegible]

RESULT	13			
LOCUS	AR042412			
DEFINITION	AR042412	11703 bp	DNA	linear
ACCESSION	AR042412			
VERSION	AR042412.1	GI:5962508		
KEYWORDS	.			
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 11703)			
TITLE	Johnston,R.E., Davis,N.L. and Simpson,D.A.			
JOURNAL	System for the in vivo delivery and expression of heterologous			
FEATURES	genes in the bone marrow			
source	Patent: US 5811407-A 8 22-SEP-1998;			
	Location/Qualifiers			
	1..11703			
	/organism="Unknown"			
BASE COUNT	3306 a	3051 c	2913 g	2433 t
ORIGIN				

Qy	1120	CAAGAAGCATTAGCTATTTTGACGACACAGATTAAGATGACGGCGTCCATTCGCTTTACAGA	1119
Dp	6283	CGAAAAACATGAGATTAAGACCCCCCAATATCCGCACTGGCTTTCAATCAGCGATGACA	6342
Qy	1180	ATACATTACAAATATGATTTGGCTGCACACTCTAATAAGAAATTGCACGTTAACCCAAATGC	1239
Dp	6343	ACACGTTACAAATATGCTCATTTGGCCGCACTAABAGAAATTGGAACTGACCGACGATGC	6402
Qy	1240	GAGAAATTACTGTCTTATGATTTGGCGGCAATTTAATGTTATGTTTCAAGAAATACCAT	1299
Dp	6403	GTCGAACGCCAACACTGACACTACGCCGCAATTCATATGCAATGCTTTGCAAAATATTCAT	6462
Qy	1300	GCATATGATGATCTGCGGATACCTTTGGCATAAACCTATTCGGCTAACACAGAAACG	1359
Dp	6463	GTAATTCGAGTATTTGGAGAGAGTTCCCTCGAAGCCAAATAGATTAACACTGAGTTTG	6522
Qy	1360	TTACGCAATATGTGACCAAACTGAAAGGGCCGAAGCAGCAGCATGTTTTCGAAATCTC	1419
Dp	6523	TCACCGCATATTTGATAGCTAGACTGAAGGCCCTTAAGCCGCGGCACTATTTTGCAAAAGCGT	6582
Qy	1420	ATTAATCTAAAAACGTTGCAGAGATACCAATGCAATTCGATCATTTGCTCATGATCTAAGAG	1479
Dp	6583	ATAATTTGGTCCCATTTGCAAGAAAGTGCCTATGATATGATTTGTCATGATGAATGAAAGAG	6642
Qy	1480	ATGCAAAAGTATCTCCCGGACGAAACATACAGAGAGCGCGCTTAAGTGCAGTTATTC	1539
Dp	6643	ACGTGAAGTTTACACACGAGCAGAAACACAGAAAGMAACCGMAAGTACCAAGTGTATC	6702
Qy	1540	AGCGTCAGATTCGCCCTTGTACGCGCTTACCTTTTGGGGATTCATCCGGAATTAATGTCGTA	1599
Dp	6703	AAGCCGAGAACCCCTTGCGCACTGCTTACTTAATTCGGGATTCACCGGGAATTAATGTCGTA	6762
Qy	1600	GACTGAATGCGGTGCTTCTGCCAAATATCCATACTCTTTTCGACATGTCAGCGGAAGTT	1659
Dp	6763	GGCTTACGGCCGCTTCTTCCAAACATTCACACGCTTTTGTGACATGTCGGCGAGGATT	6822
Qy	1660	TTGATGCGATTATTTGCTGAACATTTCCACACGCGCACCCAGATTTGGAAACGACATCG	1719
Dp	6823	TTGATGCAATCATATGACGAACACTTCAAGGAAGCGCACCGGATCTGGAAACGATATTCG	6882
Qy	1720	CGTGGTTGATTTAAAGGGAAGAGAGACGCTATTCGCAATTTGGCGGTGATATCTTAGG	1779
Dp	6883	CATCATTTGCAAAAGCCAAAGACGACGCTATGCGCTTAACCGGTCTGATATCTTGGAGG	6942
Qy	1780	ACTTAAAGTGTGACCAACCGCTCTTAAATTTGATAGAGGCGGCTTCGCAATATCAAT	1839
Dp	6943	ACCTGGGTGATGATCAACCACTACTGCACTTGATTCGAGTGGCGCTTTGGAAATATCAT	7002
Qy	1840	CTGTGACACTTACTACAGAACGAGTTTAAATTTGGTGGCATGATGAATTCGGATATGT	1899
Dp	7003	CCACCCCATCTTACCTACCGGATACTCGTTTAAATTCGGGGGAGATGATAAATCCGGAATGT	7062
Qy	1900	TCCTTAAGCGTGTGTGTCAACACTAGTGAATTCATGATTTGGTAGAGATCTACTAGT	1959
Dp	7063	TCCTCACACTTTTGTATACACAGTTTGAATGTCGTTATTCGGACAGAGATCTAAG	7122
Qy	1960	AAGCGTTAACACAGCTCAGCGTGCGGGCTCTATTCGGCGACGATTAACATATGATGATGTG	2019
Dp	7123	AGCGGCTTAAAGTCCAGATGTGCAAGCGTTCAATTTGGCGACGACATCATACATGAGAG	7182
Qy	2020	TCGCTTCCGACACCTTGATGGCGAGAGATGCGCCACTTGGCTGAACATGAAATGAAAA	2079
Dp	7183	TAGATATCTGCAAAAGAAATGGCTGAGAGGTGGCCCACTGGCTCAACATGAGAGTTAAGA	7242
Qy	2080	TTATATGATGAGATTATTTGATATCAAAAGCAACCTTATCTGTGGGGATTTATCTCTGTGG	2139
Dp	7243	TCATTCGACGATCATCTGATGAGAGACCACTTATCTTGCGGCGGAGTTATCTTTCGAG	7302
Qy	2140	ACACAGATTAACAGCACACGCTGACAGAGTCCGACACCTCTAATAAAGCTTTTAAAGCTTG	2199
Dp	7303	ATTTCGTTACTTTCAGAGAGCTGCGCGGTGGCGGACCCCTCGAATAAGGCTGTTAAATGTGG	7362
Qy	2200	GAAAAACCATTTGCCAGTGCATGATATCCCAAGACTGCGACGCGCGGCGACTGCATGATGT	2259

Db 7363 GTAAACCGCTCCGACGCGAAGACGAGCAAGACGAAAGACGCGCTCTGCTAGATG 7422
 Qy 2260 AAGCAATGCGATGGAACAGATTGGAAATTCAGACAGATTGTAAGAGCCGTAGAAATCA 2319
 Db 7423 AAAAAGGCGCTGTTTAAAGATAGTAAACAGGACATTTAGCAGTGGCCCTGACGACCC 7482
 Qy 2220 GATACGATCATCTGCGACGCGTGAATCATCAGTCTCTGTCCACGCTTACCGGAAAGCG 2379
 Db 7483 GGTATGAGTGAACAATTTACCTGTCTTACTGCGACTTGAACCTTTTGCCTCCAGACCA 7542
 Qy 2280 TTAAGAACTTCAGAGCATTAAGAGGAGCCCAATCACCTCTAGCGCTGACCTTAATAGG 2439
 Db 7543 AAAAGCATTCAGAGCCATCAGAGGGGAAATTAAGCAATCTCTAGGTTGCTTAAATAGT 7602
 Qy 2440 TGACGTGTAGACACGCACTAC-----CCACGGGCAAAATGTTTCCATACC 2486
 Db 7603 CAGCATGTACATTTCACTCTACTAATTAATCAACACCAACCAATGATGAGGATTTCT 7662
 Qy 2487 CTCAGCTGAATTTCCACCAATTTAACCTTAACCAATCCGATGGCTTAACGAGATCCAAAC 2546
 Db 7663 TTAACATGCTGCGCGCGCCCTTCCGCGCCCACTGCGATGGAAGCGCGAGAA 7722
 Qy 2547 CTCCTAGGCGCGCTGAGGCGGTTTGGCCCGCTGGCTGCTCAATCGAATCTTA 2606
 Db 7723 GAGAGGAGCGGCGCCCATGCTGCGCCCAAGGGCTGGCTTCTCAATCCAGCACTGA 7782
 Qy 2607 GGAGGTGATAGTAACTTGAATTTCAAAACAAGTACCTTAATCCCGCGCAGGTCCA- 2665
 Db 7783 CCACAGCGGTACAGTCCCTACTACTGTGACAGGCACTGACCTCAACCCCGACGTCCAC 7842
 Qy 2666 -----CCGCCAAAGAGAGAGAGTGTCTCTAAGCCCAAACTTACGCTTAAGAA 2720
 Db 7843 GCCCGCCACCGCGCCAGAGAGAGAGCGCGCCCAAGCAACACGAGCCGAGAAACCA 7902
 Qy 2721 AGAAGCAGCAAGCCAGAGAGAGCAAAAC---GCAAGCTTAACAGAGAAACGACGTA 2777
 Db 7903 AAAAGCAGAGAGAGAGAGAGAGCACTGCAAAACCAAAACCGGAAAGAGACGCGCA 7962
 Qy 2778 TGTGTATGAGTGGAGTGGACAAAGACATTTCCGATCATG---CTGAACGGCCAGTGA 2834
 Db 7963 TGGACATTAAGTTGAGGCGCCACAGATTTGTCAGATGTCAGAGAGAGAGAGATGCA 8022
 Qy 2835 ATGATATGCTGCTGGTGTGAGAGAGGCTGATGAACCACTCAGCTTGAAGGAA 2894
 Db 8023 TCGGCGACGCACTGCGCATGAGAGAGAGTAAAGTAACTCTGACGTGAAGGAAACCA 8082
 Qy 2895 TTGATTAATGAGCAATTAGCGCGCTGAAATTTGAAGAGGCTAGCATGTACGACTTGGAGT 2954
 Db 8083 TCGACACCCGTGTCTATCAAAAGCTCAAAATTTACCAAGTGTCAAGCATAGCAATGGAGT 8142
 Qy 2955 ACGCGCAGCTTCCCGAGAACATGAAATCAGACAGCTGCAATACCGACCAACCCAC 3014
 Db 8143 TCGCACAGTTGCCAGTCAACATGAGAGTGAAGTGAAGCATTCACCTACCAAGTAAACCCCG 8202
 Qy 3015 CGGGCTTCTAACAATGCGACCAAGCGCGAGTCAAGTATGAGAAATGGGAGATTTACCGGAC 3074
 Db 8203 AAGGATTTATATCTGGACCAAGAGCGGTGAGTATAGTGAAGTATGATTTACATCC 8262
 Qy 3075 CGAGAGAGTGGCGCGGAAAGCGCAACAGCGGAAAGCCGATCTGGAACAACAGAGGCAAG 3134
 Db 8263 CTCGCGGAGTGAAGGAGAGAGAGAGAGAGCGGTGCTCGATCATGATTAACCTCCGGTGGG 8322
 Qy 3135 TTGTGGCTATTGTTCTAGAGAGTGAATGAGGCAAGCGGTACGCGCTTCAAGTGTCA 3194
 Db 8323 TTGTGCGGATGTCTCTGCTGAGCTGATGAGAGAACAGAACTGCCCTTTTCGGTGTCA 8382
 Qy 3195 CTGGAACCAAGAGGGGTGACATTAAGGATACCCCGAAGTTCGAAACGGTGT--- 3251
 Db 8383 CTTGAATATGTAAGAGAGAGCAATTAAGACGACCCCGAAGGAGCAAGAGAGTGGTCCG 8442
 Qy 3252 -----CACTAGTTACAGCGCTATGCGTCTTGAATGTCAAGTTCCTCATGCAAAAC 3305

Db 8443 CAGCACCACTGTGTACCGGCAATGTTTGTCTCGGAAATGAGCTTCCATGCGACCGCC 8502
 Qy 3306 CACCCGTGTCTATTCACTGACCGCAAGAAACATCGACAGCTGTGAAGAGACGTCG 3365
 Db 8503 CGCCACATCTATACCCGGAACCTTTCAGAGCCCTGCAATCTTGAAGAGAACGTGA 8562
 Qy 3366 ACAATCCAAATTTAGACACCGCTGTGAGAACGCTTGAATGTCTATCAC---GCCGCG 3422
 Db 8563 ACCATGAGGCTTACGATACCTGTGCAATGTCATATTTGGGTGGGATCGTCTGGAGAA 8622
 Qy 3423 CCAAGCAAGATTAACCGATGACTTCACTGACCAAGTCCCTTACCTGGGCTTCTGCCCT 3482
 Db 8623 GCAAAAGAACGTTCACGAGACTTACCTCTGACCAACCCCTTACTTGGGCAACATGCTCGT 8682
 Qy 3483 ATTGCAACACTCAACCGCGTGTTCAGCCCAATTAATAATGGAAGTGGGAGCGAAT 3542
 Db 8683 ACTGCCACATCTGAACCGTGTCTGACCTCTGTTAAGATGAGACAGAGTCTGGGACGAG 8742
 Qy 3543 CTGATGATGATCGATGAGATCCAGGCTCTGGCACAAATTCGGCTACATCAGGACGCA 3602
 Db 8743 CGAGCATTAACCATACGATACGATACGACTTCCGCCAGTTTGAATAGACCAAGCGGAG 8802
 Qy 3603 CTGCGATGTACCAAAATTCGTTTACATGTCTTTGACACAGACCATGACATCAAGAG 3662
 Db 8803 CAGCAAGCGCAACCAAGTACCGCTACATGTCTGAGCAGGATCAACCGTTAAAGAG 8862
 Qy 3663 ACAATATGAGAAATATGCTATCAGCACATCTGAGCCCTGCGCTTGGCCACAAAG 3722
 Db 8863 GCACATGATGATGACATGAAGATTAAGACCTCAGAGCCGTGTGAAGGCTTATGACAAAG 8922
 Qy 3723 GGTACTCTCTGTAGCTCAATGTCTCTCAGAGTGACAGTGTAACTGATACGAGCG 3782
 Db 8923 GATCTCTTCTCTCGAAATGCTCTCAGAGGACAGCGTAAAGCTTATGATGATGATG 8982
 Qy 3783 GAGCATGTGAATTCATGACACGCTGAGAAAAGATCAGAGAGAGTGTGCTGATGAG 3842
 Db 8983 GCAACTCAGCAACGTCATGTATCACTGCGCCGCAAGATTAACCAAAATTCGTGGAGCGG 9042
 Qy 3843 AGAGATCTTGTCCCAACCGCTCCATGGAAGCTGTGAAGTGCACGTTTACGATCACT 3902
 Db 9043 AAAAATATGATCTACCTCCGTTTACCGTAAATAATTCCTTGACAGTGTACGACGTC 9102
 Qy 3903 TGAAGAGAGCTGCGCGGTTACATACATCAGTACAGGCGCCACAGCGTATAGT 3962
 Db 9103 TGAAGAAACATCTGAGGCTACATCTATGACAGGCGCGGACCGCATATATCAT 9162
 Qy 3963 CTTATCTGAGAGAGCGTCAAGGAGTATCAATTAACCACTTCTGCGAAGACGTC 4022
 Db 9163 CTTACTGGAAGATCATCTAGGGAAGATTTACGCAAGCCGCACTGAGGAAGAACATTA 9222
 Qy 4023 CTTACGAATGTAAGTGTGGGACTACAGCACAGGTATCGTGAAGCGGAAACGAAGTGA 4082
 Db 9223 GGTATGAGTGCAGTGTGGGAGCTACAAAGCCGGAACCGTTTGAACCGGACCGAAATCA 9282
 Qy 4083 ACGGCTCATTAAGCAAAACAGTGAATGCTTACCAAGAGCGAACCAAGAAATGGTCT 4142
 Db 9283 CTGCTTGCACCGCATCAAGCAGTGGCTCGCTTATGAAGAGACCAACGAAGTGGTCT 9342
 Qy 4143 TCAACTGCGCGATCTTATTAAGCAACAGACCACTGATGCAAGTGAATTTGCACTTC 4202
 Db 9343 TCAATCACCGGACTTATGATGACATGACACCAAGGCGCCAAAGGAAATTTGATTTGC 9402
 Qy 4203 CATTCGCTTGAACCGACAGTCTGCGCGGTTCCGTTAAGTCTACAGCCCTACAGTCA 4262
 Db 9403 CTTTCAAGTTGATCTCGAGTACCTGCAAGTGTCTCTGTGTCACAGCGCGCAATTAATAC 9462
 Qy 4263 AGTGTCTCAAGGATCACTCCACTGACTGATGACGACCAATTTGCTGCAACGA 4322
 Db 9463 ATGCTTTAAACATATAGCTCTCAATTTAGATACAGACCACTTGAACATTCCTACACCA 9522
 Qy 4323 GAAATATGGGCGCTGCGAGACGCAACAGCAGATGATTAACAGGTTTACATCTCAGGA 4382
 Db 9523 GGAAGCTAGGGGCAAAACCGGAACCAACCACTGAATGATGTGTGGAAGAGCGGTCA 9582

QY	4383	ATTTTTCGTGAGGGGCGAAGAGGGCTGGAGATACGATAGGGGTAACTATGAACCACTGATAGAG	4442
Db	9583	ACTTCAACCGTGCACCGAGATGGCTGTGAATATCATATGAGGAAATATATGAGCCAGTGAAGG	9642
QY	4443	TCTGGGCGCCAGAGATCGGCACACGAGCACCACTGATGAGCGGCAATGATCATATCC	4502
Db	9643	TCTATGCCCAAGATCAGACACACAGAAACCTCTACGGATGGCCACACGAATATGTCACAGC	9702
QY	4503	ACTATATCATCGGCATTCACGCTTACACTGTCACTTGTGCTGTGTGCTGCTCTTGTCTA	4562
Db	9703	ATTACTACATCGCCATCTCTGTGTACACCAATTTAGCCGTCGATCAGATCAACGCTGGCA	9762
QY	4563	TCTGTGTAGGCACTGTGCATCATCAGACGCTTGATCGCCAAAGCAAGAGACTGCTTA	4622
Db	9763	TGATGATTTGGGGTAAACGGTTGCAGTGTATATGTGCTGTAAAGGCGCCGTCGATGCTGA	9822
QY	4623	CGCCATACGCGCTTGCACCCGAAGCGCAACGGATCCACAGATTAAGAGGTTTGTGCTGA	4682
Db	9823	CGCCATACGCGCTTGCACCCGAAGCGGTAATCCCACTTGTGCTGGACTCTTGTGCTGG	9882
QY	4683	TTGCGCCAACCAACGCTGAAACATTTGAGAAATTTGAAACCATCTGTGTTTAAACACC	4742
Db	9883	TTAGGTGGCCCAATGTGTGAACGTTTCAACGAGACATAGATTACTGTGTGGAACAGTC	9942
QY	4743	AACCGTTTCTGTGGGCACAGTTGTGCATTCCTCTGCGACGCGCTTGTATCTGTTCGCT	4802
Db	9943	AGCGGTTCTTGTGGGTCCAGTTGTGCATPACTTTGGCGGCTTTCATCGTTTCAATCCGCT	10002
QY	4803	GCTTTATATGCTGATGCTTTTATTTGGTGTAGGCGCTGTGCGGGGAAGGTAGAG	4862
Db	10003	GCTGCTCTGTGCTGCTGCTCTTTTATGTGTGTGCGGCGCTACCTGTGGGAAGGTAGAG	10062
QY	4863	CCTTGCAACATCGACCACTGTGCGCAATGTTCGCGGGATCCCGTAAAGGCGTTGTGTG	4922
Db	10063	CCTAGGAACATGCGACCACTGTTCATATGTGCGACAGATACCGTAAAGGCACTTGTGTG	10122
QY	4923	AACGCGCAGGTTACGCGGCACCTTAACCTGTGAGATCAGCGTGTCTATCGGAATTAAC	4982
Db	10123	AAAGGGCAGGGTATACCCCGCTCAATTTGGAGATCATGTCATGTCTCGAGAGTTTGC	10182
QY	4983	CTTCACTTAAAGAGTATGCTGACTCTGGCAATTTCCACATGTCATTTCTTACCCACAG	5042
Db	10183	CTTCCACCAACCAAGATGATCATTCCTGCAAAATTCACATGTGTGTCCCTCCCAAAA	10242
QY	5043	TTAATGCTGCGGGTCCCTCGAGTGCAGAGCATCTCAAAAGGCGGATTAACAATGCCG	5102
Db	10243	TCAATGTGTGGCTCTCTTGTGAATGTAGCCGCGCTCATGTGAGACTATCCTGTCAAG	10302
QY	5103	TTTTTGGCGGTGTACCTTTTCAATGTGGGAGGCGCACATGCTCTGTGTGACGTGAGA	5162
Db	10303	TCTTTCGGAGGGGTCTACCCCTTATATGTGTGGGAGGAGCGCAATGTTTTTGGACAGTGA	10362
QY	5163	ACACACACATGATGAGAGCGGTATCGTGAATTCGTCACACTCATATGATCAGCAG	5222
Db	10363	ACACCCAGATGAGTGAAGGCGTACCTCAACCTGTCAAGATTCGCGTCTGACCAAGCGCC	10422
QY	5223	TGCGACTTAAAGTTTACACACAGCTGCTGTGAAGTCGCGCTGGGTATATAGTAAACGCA	5282
Db	10423	AGCGGATTAAGGTTCACACTGTGCGCATGAAAGTGAAGTCTGCTATATATGTACGGGA	10482
QY	5283	CCACCGCGCACTTGATACGTTGTCTCATGTGGCGTCAAGCCAGGTTCTCATCGGACCTGA	5342
Db	10483	CTTACAGATTTCTATGATGTGTAGCTGAAGCGAGTCAACACGGAAGAGTCTAAAGACTTGA	10542
QY	5343	AGGTCTATGACAGGCGCGATATCAGCCGCTTTTCAACCTTTGACCAATAAGTGTGATCA	5402
Db	10543	AAGTCATATGCTGACCAATTTACGACTGTTTACGCAATTCGATTCATTAAGTCGTATCC	10602
QY	5403	GAAGGGGCGTTGTTTCAACTAGACATTCCTCCTAGATATGAGGTATGAACAAGAGCGCT	5462
Db	10603	ATCGGCGCTGTGTATCAACTATGACTTCCCGGAATATGTGAGGAGATGAACCAAGAGCGT	10662

OY		5463	CGGGAAATTCAGACATCCCTGCTTAAGTCAAGCATAGTAGCCCGACATGCATAC	5522
Db		10663	TTGAAGACATTCAAGTACTCCTTGATCTAGCAAGATCTCATGCCAGACACAATTA	1072
OY		5523	GCGCTGTGAAGCCCTTCTGTCAAGAACAATCCACGCCCCCTCACACCAGCAGTACAGGT	5582
Db		10723	GGCTAATCTMAACCTTCCGCCAAGAACGTGATGTCCCCTACACGACAGGCCGATCAGAT	1078
OY		5583	ATGAAATGTGGAGAACAACTCAGACGACCCCTGCAAGAAAAGACCAATTTGGATGTA	5642
Db		10783	TTGAGATGTGGAAAAAACAACTCAGGCCGCCCACTGCAGAAAAACGCACTTTGGGGGTGA	1084
OY		5643	AAATTGGAAGTGAAGCCCTCTGGAGCCGTCTAACTGTGCTTACGGGCATAATCCCTATCTCGA	5702
Db		10843	AGATTGCAAGTAAATCCCTCCGACCCGGTGAATGTTTATACGGGAACATTTCCATTTCTA	1090
OY		5703	TTGACATCCCTGATGACAGCTTTTGTGATCATCAGAAATCAACCAATTTTGAAGTTA	5762
Db		10903	TTGACATCCCGAACGCTGCCCTTATACGACATCAATGACACCACCTGCTCTCAACAGTCA	1096
OY		5763	GCTSCACAGTAGACAGACTGCATTTATTTCTGACACTTTGTGTGTTCTTAACATTAACGT	5822
Db		10963	AATGTGAAGTCAAGTAGTGCACTTATTACAGAGACTTCGGCGGGATGGCCACCCTGCAGT	1102
OY		5823	ACAAAGCTGACAGGAGGAGACATTTGTCCAGTCTACTCTCCACTCCAGACAGCTGTTTGA	5882
Db		11023	ATGTATCCGACCGGGAAGGTCAAGCCCCGCTGACATTCGCATTCGAGACAGCAACTCTCC	1108
OY		5883	AGGAAGCGACCAACATATGTGACTCCGCTAGGACAGCATPAACATTTTATAGCAGATCGA	5942
Db		11083	AAGAGTGCACATGTCATGTCTGTGAGAAAGAGAGCGGTGACAGTACATTTTACACGCGCA	1114
OY		5943	GCCCCAAGCAAAATTTTATAGTTTCGCTATGCGGCAAGAAAGTCCACCTGCATGCTGAAT	6002
Db		11143	GTCCACAGCGCAATTTATCGTATCGCTGTGTGGGAAGAAAGACATGCAATGACAGAAAT	1120
OY		6003	GTAACCAACCGGCCGACCAACATTAATGGAAGAACACATAAAGTCAGCAAGAAATTCGAG	6062
Db		11203	GTAACCAACCAAGCTGACCAATATCGTGAACAACCCCGCAAAAATGACCAAGAAATTTTCAG	1126
OY		6063	CGGCAAGTTCCAAACATCTTGGAAACGTGGCTGTGACATGTTTGGGGGGAGCATCATCC	6122
Db		11263	CGCCATCTCAAAAACATCATGAGTTGGCTGTGCTTCCCTTTTGGGGGGCGCTCGTGC	1132
OY		6123	TCAATTTGTAGAGCTTATAGTGTGTGCTGCAGCTCTATGCTTATAAACACAGTAGAT	6182
Db		11323	TATTAATTAAGACCTTATGATTTTTTCTGTGACAGATGATGTGACTAGCAACAGAAAT	1138
OY		6183	GACTGAGCGCGGACACTGA-CATAGCCGTAAACCTCGATGTACTTCGAGAGAAAGCTGT	6241
Db		11383	GACCGCTACGCCCCCAATGATCCGACGCAAAACTCGATGTACTTCGAGAGAACTGATGT	1144
OY		6242	GCATPATGCCACGCGCGCT 6261	
Db		11443	GCATPATGCATCAGGCTGT 11462	
<hr/>				
RESULT 14				
LOCUS	AR096558	11703 bp	DNA	linear PAT 08-SEP-2000
DEFINITION	Sequence 8 from patent US 6008035.			
VERSION	AR096558			
KEYWORDS	AR096558.1 GI:10025468			
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclashed.			
AUTHORS	1 (bases 1 to 11703)			
TITLE	Johnston,R.E., Davis,N.L. and Simpson,D.A.			
JOURNAL	System for the in vivo delivery and expression of heterologous genes in the bone marrow			
FEATURES	Patent: US 6008035-A 8 28-DEC-1999;			
	Location/Qualifiers			

source 1. .11703
/organism="unknown"
BASE COUNT 3306 a 3051 c 2913 g 2433 t
ORIGIN
Query Match 36.8%; Score 2388.4; DB 6; Length 11703;
Best Local Similarity 64.7%; Pred. No. 0;
Matches 3698; Conservative 0; Mismatches 1981; Indels 41; Gaps 8;
QY 580 ACTGACGGTATGAAGCGGAGCGGTATTTTCTCATGGGAAACGGCCAGGTACCTTC 639
DB 5746 ACTGACCTAACCGGGGTGGTGGTACATTTTTCAGCGACACAGCGCCCTGGGACCTTC 5805
QY 640 AACGAAATCAAGTACGTCAATGTAACTACAAAGAACCTATTTGGATGGGCGCTTCATG 699
DB 5806 AAAAGAAAGTCCGTTCTGCGAAGACCAAGCTTACAGAACCGACCTTGGAGCGCAATCTTCG 5885
QY 700 AGAAGTATACGCGCGCGCTCGATCTCGAAGAGAGAAATGTACAGAAAGAACTGC 759
DB 5866 AAAGAAATTCATGCGCGGTGCTCGACAGTCCGAAAGAGAACACTCAACTCAGGTACC 5925
QY 760 AATTATCGGCTCTGAAGGAAATGAAGCAGTATCATACGAAAGTAAATATGA 819
DB 5926 AGATGATGCCCAACGAAAGCAAAAGTAGTACCACTCTCGTAAAGTAAATATGAGA 5985
QY 820 AAGCAATTAACGCGGAGCACTCTTCTGGATTGGGCACTATCTATCATCAGAGTGA 879
DB 5986 AAGCATATACCACTGAGCGCACTACTGTCAGACTACACTGTATTAATCTTCCACAG--- 6042
QY 880 ATCTCTGAGGTGTACAGATCAATTAATCTGTACCAATCTACTCGTCAACGGTAAATTA 939
DB 6043 ATCAGCCAGATGTATTAAGATCACTACGAAACCACTGTACTCCAGTAGCGTACC 6102
QY 940 ACAGGTTTACATCTGCAAGGTCCGCGTTAAAGTGAATCTTATCCAAAGAAATT 999
DB 6103 CGAACTCTCGATCCACAGATTCCTGTGCTGTACCACTATCTGTGATAGAACT 6162
QY 1000 ACCCTACAGTAGCAGGTATATGTATTAACAGATGAATACGATCGTATCTTACATGGTGG 1059
DB 6163 ATCCGACAGTACGATCTTATACAGATTAAGTACAGATGATGCTTACTTGGATATGGTAG 6222
QY 1060 ACGGCGCATCGTCTGTCTAGATACAGCACTTTTTCGCGCTAACTGAAAGCTAAC 1119
DB 6223 ACGGAGCAGTCCGCTGCTGATCTGCACTTCTGCGCGCTAAGCTTGAAGTTAAC 6282
QY 1120 CAAGAAGCATAGCTATTTTGGAGCGAAGATTAAGATCAGCCGTCCCATCGCTTATACGA 1179
DB 6283 CGAAAAAATGAGATTAAGACCCCGAATATCCGAGTGCCTTCATCAGCGATGACGA 6342
QY 1180 ATACATTAACAAATGTATTTGGCTGAGCTACTTAAAGGAATTTGCAACGTTACCCAAATGC 1239
DB 6343 ACAGCTATCAAAATGTGCTACTTTCGCCCACTTAAAGAAATTTGCAAGCTCAGAGATGC 6402
QY 1240 GAGAAATTAAGCTTTAGATTTGGCGCGATTTAAATGTTGATTTTCAAGAAATACGAT 1299
DB 6403 GTGAAGTCCCAACACTGGAATCAGCACTTCAATGTGATGCTTTCGAAATATATGAT 6462
QY 1300 GCATGATGAGTACTGGGATACCTTTGGCGATTAACCTTATTCGGCTAACTACAGAAAG 1359
DB 6463 GTAATGACGAGTATTTGGAGAGATTCCTCGAAGCCAAATTGATTAACAATGATTTG 6522
QY 1360 TTACGCAATATGTGACAAAGCTGAAAGGGCGGAAAGCAGACATTTGTTTTCGAATCTC 1419
DB 6523 TCACCGCATATGTAGCTAGCTGAAAGGCCCTTAAAGGCCCGCGCACTATTTTTCAAAGAGCT 6582
QY 1420 ATAAATCTAAACCGTTGACAGAGATACCAATGATCAATTCGTATGATGATTAAGAGAG 1479
DB 6583 ATAAATTTGTCCTCATTTGCAAGAGTGCCTATGATGATGATTCGTATGATGATGAAGAG 6642
QY 1480 ATGTCAAGTATCTCCGCGCAGAACTATACAGAGAGCGGCTTAAGGTGAGGTATATTC 1539
DB 6643 ACGTGAAGATTAACACAGGACGAAACACACAGAAAGAAAGACGAAAGTATCAAGTATATC 6702

QY 1540 AGGCTGAGATCCCTTGTACTACCGCTTACCTTTGCGGATATCATTCGGAATTAATGCCGA 1599
DB 6703 AAGCGGAGAACCCCTGCGACATGCTTACTATGCGGGAATTCACCGGGAATTAATGCCGA 6762
QY 1600 GACTGAATGCGGTCTTCTGCGCAATATCAATCTCTTTCGACATCTCAGCGGAATTT 1659
DB 6763 GGGCTTACGCGCGCTTGTCTTCCAAACATTCACACGCTTTTGTGACATCTCGCGGAGATT 6822
QY 1660 TTGATGCAATTAATTTGCGAATTTTCCACAGCGCGACCCAGATTTGGAACCGACATTC 1719
DB 6823 TTGATGCAATTAATTTGCGAATTTTCCAGAGGCGACCCGATCTGGAAGCGATATTC 6882
QY 1720 CGTCTGTTGATTAAGGGAAGCAGACGCTATTCGCAATTTGCGCGTGTATGATCTTTCAGG 1779
DB 6883 CATCATTTGCAAAAGCAAGCAGACGCTATTTGCGTTAACCGGTGTATGATCTTTCAGG 6942
QY 1780 ACTTATGTTGACCAACCGCTCTTATGATTTGATAGGCGCGGCTTGGCAATATCAT 1839
DB 6943 ACCTGGGTGTGATTCACCACTACTGCACTTGTATCGAGTCCGCTTTGGAGAAATATCAT 7002
QY 1840 CTGTGACCTTACCTACGAAAGCAGAGGTTTAAATTTGGTGCATGATGAATTCGGTATGT 1899
DB 7003 CCACCCATCTACCTACGCGGTACTGTTTAAATTCGGGCGATGATGAATTCGGAATGT 7062
QY 1900 TCTTAACGCTGTTTGTCAACACACTAGTCAATATCATGATTTGCTAGCAGATTAACGTG 1959
DB 7063 TCTCACTACTTTTGTCAACACAGTTTGAATGTGTTATCGGACAGAGTACTTAAG 7122
QY 1960 AACGTTTAAACAGTCAAGCGTGGCGGCTCTATCGCGACATTAATATGTCATGTG 2019
DB 7123 AGCGGCTTAAAGGTCAGATGTGCAAGCTTTCATGCGACGACATCATATCATATGAG 7182
QY 2020 TCGTCTCCGACACTTGAATGCGGAGAGATGCGCACTTGTGGTGAATGGAATGAATAA 2079
DB 7183 TAGATTTGAACAAAGAAATGCTGAGAGTGGCGCACCTGCTCAATGAGGTAAAGA 7242
QY 2080 TTATGATGAGTATTTGTATGATCAAAACACTTCTGTGGGGAATTAATCCTGTGTC 2139
DB 7243 TCATGACGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7302
QY 2140 ACCAGATTAACAGGACAGCTCTGCAAGATGCGAAGCTCTTAAAGGCTTTTAAAGCTTG 2199
DB 7303 ATTCGTTTACTTTCACAGCGTGGCGCTGCGGACCCCTGAAAGGCTGTTTAAAGTTGG 7362
QY 2200 GAAACCATTTGCGATGATGATACCCAGACTGCGACCGCGCGGCGCATGATGATG 2259
DB 7363 GTAAACCGCTCCAGCGGACGAGCAGACGAGACGAGACGAGACGAGACGAGACGAGATG 7422
QY 2260 AAGCAATGCGATGGAACAGAAATGGAATTAACGAGATTAAGTAAAGCGCGTAAATCCA 2319
DB 7423 AAACAAAGGGGTGTTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 7482
QY 2320 GATTCGATCATATCTGCGAGGCTGATCATCATGCTTCTGTCCAGCTTACCGGAAGCG 2379
DB 7483 GGTATGAGGTAGCAATATTAACATCTGCTCACTGGAATGAGAACTTTTTCGCCAGGCA 7542
QY 2380 TTAAAGACTTCAAGAGCATTAAGAGGAGGCCCAATCACTCTTACAGCTGACCTTAATAGG 2439
DB 7543 AAAGAGATTTCCAGGCAATGAGGAGGGAATTAAGCATCTCTACGAGTGTCTTAATAGT 7602
QY 2440 TGAAGTATGAGACGCACTTAC-----CCACCGGAGAAATGTTTCCATACC 2486
DB 7603 CAGCATATGATTAATTCATGATGATTAATTAACACACACACACATGATTAAGAGATTC 7662
QY 2487 CTCAGCTGAACCTTTCACAGCTTATACCTTCAAAATCCGATGCTTACCGGATCAAAAC 2546
DB 7663 TTAAATATGCTCGCGCGCGCGCTTCCCGGCGCCCACTGCAATGTGAGGCGCGGAGAA 7722
QY 2547 CTCCTAGCGCGCGTGGAGCGGCTTGGCGCGCGCGCTGAGCTCAATGGAAGATCTTA 2606
DB 7723 GAGGCAAGCG 7782
QY 2607 GAGGTGATATGATCACTTGAATTTCAACAGATCACTTAATCCCGCGCGAGTCCA- 2665

Db 7783 CCACAGCCGTAGTCCCTAGTCATTTGGACAGGCACTGACCTCAACCCCACTGAC 7842
Qy 2666 -----CCGCCAAGAGAGAGAGAGTGTCTCTAACCCTAACCCTAAGTAAAG 2720
Db 7843 GCCCGCACCCGCCAGAGAGAGAGGCGCCCAACCAACACCGAAGCGAAGAAACCA 7902
Qy 2721 AGAAGCAGCAAGCCAAAGAGAGAGAAAC---GCAGCCTTAAACAGAGAGAAACGACACGTA 2777
Db 7903 AAAGCAG 7962
Qy 2778 TGTGTATGAGTGTGAGTGTGAGCAAGACATTTCCGATCATG---CTGAAACGCGCAAGTGA 2834
Db 7963 TGGCATTAACTTGGAGGCGCCGACATGTTGTCAGTCAAGAGACGAGAGAGATGTC 8022
Qy 2835 ATGATATGCTGCTGCTTCCGAGAGAGGCTGATTAACCACTCCAGCTTGAAGAGAAA 2894
Db 8023 TCGGCGACGACATGCGCATGAGAGAGAGATATGAACTCTTGACGCTGAAGAGAGCA 8082
Qy 2895 TTGATATGAGCAATTTAGCGGCGCTGAATTTGAAGAGAGAGAGAGAGAGAGAGAG 2954
Db 8083 TCGACACCTGTGTGTATCAAAAGCTCAATTTACCAAGTCTGACAGATACAGATGAGT 8142
Qy 2955 ACGGCGACGTTTCCCGAAGCATGAATCAGACACGCTGCAAGTACACAGAGCAAAACAC 3014
Db 8143 TCGCAGATGCGACATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8202
Qy 3015 CGGCTTTTACACATGCGACACACGCGGAGTCCAGATGAGAGAGAGAGAGAGAGAGAG 3074
Db 8203 AAGGATTTATTAATCGGAGACACAGAGGCGTGCATATGAGAGAGATTTACATCC 8262
Qy 3075 CGAGAGAGTGGGCGGAG 3134
Db 8263 CTCGCGAGATGAG 8322
Qy 3135 TTGTGCTATTGTTCTAGAGAGTGCATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3194
Db 8323 TTGTGCGATGATGCTCGGTGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8382
Qy 3195 CTTGGAACCAAG 3251
Db 8383 CTTGGAATGATTAAG 8442
Qy 3252 -----CACTAGTTACAGCGCTATGCTGTGATGTCAGTTCATGCGCAAG 3305
Db 8443 CAGACACCTGTGTCTCGGCAATGTGTGTCTGAGAGAGAGAGAGAGAGAGAGAGAG 8502
Qy 3306 CACCGGTGTGATTTCACTGACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3365
Db 8503 CGCCACATGCTATACCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8562
Qy 3366 ACAATCCAAATTTAGACAGCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3422
Db 8563 ACCATGAGCGCTACCATACCTGCTCAATGCGATTTGCGGTGCGAGATGCTGCGAGAA 8622
Qy 3423 CCAAGCAGAGATTAACGATGATCTTCACTGACAGAGAGAGAGAGAGAGAGAGAGAGAG 3482
Db 8623 GCAAAAGAGGCTACGACGATCTTACCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8682
Qy 3483 ATTGACAGACTCAACGCGGTGTTTCAAGCCCAATTTAAATTTAGAGAGAGAGAGAGAG 3542
Db 8683 ACTGACACATTAAGTGAACCGTGTCTTCAAGCCCTGTTAAATGAGAGAGAGAGAGAGAG 8742
Qy 3543 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3602
Db 8743 CGGAGCATTAAGACCTTAAGCATGACAGATCTCGCCAGATTTGAGATACAGCAAGAGAG 8802
Qy 3603 CTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3662
Db 8803 CAGCAGAGCAAAAG 8862
Qy 3663 ACAGATGAG 3722
Db 3722

Db 8863 GCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8922
Qy 3723 GGTACTTCTGTTAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3782
Db 8923 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8982
Qy 3783 GAGATCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3842
Db 8983 GCAACTGAGCAAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9042
Qy 3843 AGAGATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3902
Db 9043 AAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9102
Qy 3903 TGAAGAGAGAGTGTGCGGAGTCAATGATGATGATGATGATGATGATGATGATGATGATG 3962
Db 9103 TGAAGAGAGAGAGTGTGCGGAGTCAATGATGATGATGATGATGATGATGATGATGATGATG 9162
Qy 3963 CCTATCTGAGAGAGAGTGTGCGGAGTCAATGATGATGATGATGATGATGATGATGATGATG 4022
Db 9163 CCTATCTGAGAGAGAGTGTGCGGAGTCAATGATGATGATGATGATGATGATGATGATGATG 9222
Qy 4023 CCTATCTGAGAGAGAGTGTGCGGAGTCAATGATGATGATGATGATGATGATGATGATGATG 4082
Db 9223 CGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9282
Qy 4083 ACGGCTGATCTTAAG 4142
Db 9283 CTGTTGAGACCGCATCAAG 9342
Qy 4143 TCAACTGCGCGGATCTTATTTAGGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4202
Db 9343 TCAACTGCGCGGATCTTATTTAGGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9402
Qy 4203 CATTCGCTTGAACCGACAGTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTG 4262
Db 9403 CTTTCAAGTGTATCCCGAGTACTGATGATGATGATGATGATGATGATGATGATGATGATG 9462
Qy 4263 AGTGTGTTCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4322
Db 9463 ATGCTTTTAAACATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9522
Qy 4323 GAAATTTGGGCTGCGAG 4382
Db 9523 GGAAGATGAG 9582
Qy 4383 ATTTTGTGTGAG 4442
Db 9583 ACTTCAACGCTGACCGAG 9642
Qy 4443 TCTGAGCGCGAG 4502
Db 9643 TCTATGCGCGAG 9702
Qy 4503 ACTATTTATCATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4562
Db 9703 ATTACTACATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9762
Qy 4563 TCTGTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4622
Db 9763 TGATGATGAG 9822
Qy 4623 CGCCATGAGCGCTTGAACCGAG 4682
Db 9823 CGCCATGAGCGCTTGAACCGAG 9882
Qy 4683 TTGCGCAACCAAG 4742
Db 9883 TTGAGTGTGCGCAAG 9942
Qy 4743 AACGTTTCTGTGAG 4802
Db 9943 AGCGGTTCTTGTGAG 10002

QY 4803 GCTTTTCATGCTGATGCTTTTATTTATTTGTTGAGGCGTCTGCTGGGAAAGGTAGACG 4862
 DB 10003 GCTGCTCTGCTGCTGCTGCTTTTATTTAGTGTGCGGCGCTTACTGGCAAGGTATGACG 10062
 QY 4863 CTTTGCAGATGCGACCACTGTGCGAAATGTTCCGGGGATCCCGTATAGAGCGTTGTGTCG 4922
 DB 10063 CTTATGAAATGAGACCACTGTTCCAAATGTGCGACAGATACCGTATATAGGCACTTGTG 10122
 QY 4923 AAGCGGAGGTTACGCGGCACTTAACTGAGATCAGGCTGTCTCATCGGAATTAACAC 4982
 DB 10123 AAGGGGAGGATAGCCCGCTCAATTTGAGATCATGTCTATCTCCGAGGTTTTCG 10182
 QY 4983 CTTTCAATTAACAGAGATGAGCTGCAATTTCCACAGATCATTTCTTACACAGAG 5042
 DB 10183 CTTTCAATTAACAGAGATGAGCTGCAATTTCCACAGATCATTTCTTACACAGAG 10242
 QY 5043 TTAATGCTGCGGCTCCCTGAGTGCAGAGCATCTGCAAGCGGATTTACATGCGCG 5102
 DB 10243 TCAATGCTGCGGCTCCCTGAGTGCAGAGCATCTGCAAGCGGATTTACATGCGCG 10302
 QY 5103 TTTTGGCGGTGTATCCCTTTATGTTGGGAGCGCAATGCTTCTGTGACGTGAGA 5162
 DB 10303 TCTTGGAGGGGTCTACCCCTTTATGTTGGGAGCGCAATGTTTGGCAGACGTAGA 10362
 QY 5163 ACAACAATGAGTGAAGGCGGTAGCGTTCGCTGAGACTGCACTATGATGATGACGAG 5222
 DB 10363 ACAGCCAGATGAGTGAAGGCGGTAGCGTTCGCTGAGACTGCACTATGATGATGACGAG 10422
 QY 5223 TCGCACTTAAAGTTTACACAGCTGCTGAAAGTCCGCTGCTATGATATAGCGCAACA 5282
 DB 10423 AGGCGATTTAAGTGTACACTGCGCGATGAAAGTATGAGCTGCTATGATGATGAGCAACA 10482
 QY 5283 CCACCGGCGACCTGAGTATGCTTTTCAATGCGCTGACGCGGCTTCTTACGCGGACCTGA 5342
 DB 10483 CTACAGACTTTCTAGATGTGTACGTGAACGAGTCAACAGCAACGCTTAAACACTTGA 10542
 QY 5343 AGGCTATGAGAGGCGCGATATACGCGCTTTTATACCTTTGACCATTAAGGCTGCTATCA 5402
 DB 10543 AAGTCAATGCTGAGCACTTTTACAGCATGCTTTTACGCACTTGCATTAAGGCTGCTATCC 10602
 QY 5403 GAAAGGGCTTGTTTTACACTACACTTCCCTGAGTATGAGCTATGAAACAGAGAGCGT 5462
 DB 10603 ATGCGGCGCTGTGTATACAGTATGCTTCCGGAATATGAGCGATGAACAGAGAGCGT 10662
 QY 5463 TCGCGATTTTCAAGCATCTCTGCTTGTATGCTTACAGACATATGATGCGGCACTGACATAC 5522
 DB 10663 TTGGAGCATTTCAAGCTACCTCTTGACTAGCAAGATCTCATGCGCAGACAGACATTA 10722
 QY 5523 GGCTGCTGAAGCCCTTGTGTAAGAACTCCAGTCCCTTACAGCCAGAGCATGACAGGCT 5582
 DB 10723 GGCTACTCAAGCCCTTGTGTAAGAACTCCAGTCCCTTACAGCCAGAGCATGACAGGCT 10782
 QY 5583 ATGAATGTGGAAGAACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5642
 DB 10783 TTGGAGATGTGGAAGAACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 10842
 QY 5643 AAATTTGAAGTGAAGCTCTGTGAGAGCTTAACTGTGCTTACGGGCAATCCCTTATCTGGA 5702
 DB 10843 AGATTTGAAGTGAAGCTCTGTGAGAGCTTAACTGTGCTTACGGGCAATCCCTTATCTGGA 10902
 QY 5703 TTGACATCCCTGATGAGAGCTTTTGTGATCATGAGATCAGAACTCAATTTTAAAGTTA 5762
 DB 10903 TTGACATCCCTGATGAGAGCTTTTGTGATCATGAGATCAGAACTCAATTTTAAAGTTA 10962
 QY 5763 GCTGCAAGTGAAGAGCTGATTTATTTCTGAGACTTTGTGTCTTCTTAACTATAGAGT 5822
 DB 10963 AATGTGAAGTGAAGAGCTGATTTATTTCTGAGACTTTGTGTCTTCTTAACTATAGAGT 11022
 QY 5823 ACAAGCTGAG 5882
 DB 11023 ATGTATCCAG 11082

QY 5883 AGGAGCGACCAACATGATGATGCTGCTGAGAGAGATTAACATATTTTAGACATCGA 5942
 DB 11083 AAGAGTGCAGATGATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11142
 QY 5943 GCCCAGAGCAATTTATATGTTTCCGTATGCGGCAAGAGTCCACTGCAATGCTGAAT 6002
 DB 11143 GTCACAGCGGAACTTATGATGCTGCTGTGTGGAAGAGAAACATGCAATGCAAGAT 11202
 QY 6003 GTAAACAGCGCGGAG 6062
 DB 11203 GTAAACAGCGCGGAG 11262
 QY 6063 CGGAGTTTCCAAACATCTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6122
 DB 11263 CCGCATCTCAAAACATGATGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 11322
 QY 6123 TCAATTTGATGAGACTTATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6182
 DB 11323 TATTAATTTATAGACTTATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 11382
 QY 6183 GACTGAGCGGAG 6242
 DB 11383 GACTGAGCGGAG 11442
 QY 6242 GCATTAATGCGAGCGCGCT 6261
 DB 11443 GCATTAATGCGAGCGCTGCT 11462

RESULT 15
 BD071123 11703 bp DNA linear PAT 27-AUG-2002
 LOCUS System for the in vivo delivery and expression of heterologous
 DEFINITION genes in the bone marrow.
 ACCESSION BD071123
 VERSION BD071123.1 GI:22616726
 KEYWORDS JP 2001515348-A/4.
 SOURCE unidentified
 ORGANISM unidentified
 unclassified.
 REFERENCE 1 (bases 1 to 11703)
 AUTHORS Johnston,R.E., Davis,N.L. and Simpson,D.A.
 TITLE System for the in vivo delivery and expression of heterologous
 JOURNAL genes in the bone marrow
 COMMENT Patent: JP 2001515348-A 4 18-SEP-2001;
 UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL
 OS Girdwood S.A.
 PN JP 2001515348-A/4
 PD 18-SEP-2001
 PF 18-FEB-1998 JP 1998536760
 PR 19-FEB-1997 US 08/801263
 PI ROBERT E JOHNSTON,NANCY L DAVIS,DENNIS A SIMPSON PC
 C12N15/86,C12N15/33,C12N7/01,C12N5/10,A61K39/12,A61K48/00 CC
 System for the in vivo delivery and expression of heterologous
 CC genes in
 the bone marrow
 FH Key
 FT source 1.11703 Location/Qualifiers
 FEATURES
 source 1.11703 Location/Qualifiers
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 BASE COUNT 3306 a 3051 c 2913 g 2433 t
 ORIGIN
 Query Match 36.8%; Score 2388.4; DB 6; Length 11703;
 Best Local Similarity 64.7%; Pred. No. 0;
 Matches 3698; Conservative 0; Mismatches 1981; Indels 41; Gaps 8;
 QY 580 ACTGACGGTATGAAGCGGAGCGGTATTTTCTCATGGAAGAGCGCAAGGTACCTTC 639

Dh 5746 ACTGACTAACCGGGGTAGGTGGGTACATATTTTTCAGACGACACAGGCCCTGGGACTTGC 5805
Qy 640 AACAGAAATCAGTACGTCAATGTAACTAACAGAACTTATTTGATCGGGCGTCCATG 699
Db 5806 AAAAGAAATCGCTTTCGAGAACAGCTTACAGAACTTTCGAGCGCAATGTCCTGG 5865
Qy 700 AGAAGTATTACGCCCCCGGCTCGATCTTCGAAAGAGAAAATGTTTACAGAAAGAACTGC 759
Db 5866 AAAAGAAATCATGCCCGGTGCTCGACAGTCGAAAGAGAAACAACCTCAAACTCAGGTACC 5925
Qy 760 AATTATGGGCTCTTGAAGAAATAGAGCAGATTCATCAGAAAAGTAAAGAAATATAGA 819
Db 5926 AGATATGCCCAACCAAGCAACAAAGATGATCAGCTCTGTAAGATGAAATTCAGA 5985
Qy 820 AAGCAATTACAGCGGAGCACTATTTCTGATTTGGGCAATATCATCATGAGAAATGA 879
Db 5986 AAGCAATTACCACTAGAGCACTACTGTACAGGACTACGACTGTATTAACCTCTGCCACAG--- 6042
Qy 880 ATCTGTGAGTGTTCAGAGTCAATTTCTGTACCAATCTACTCTGTCAACGCTAATTA 939
Db 6043 ATCAGCAGAAATGCTATTAAGATCACCTATCCGAAACATTTGACTCCAGTACGTAACGG 6102
Qy 940 ACAGGTTTACATCTTCAGAGGTGCGCGTTAAAACTGCACTTACTTATCCAGAAAT 999
Db 6103 CGAATCTCCGATCCAGATCCAGATTCGCTGTAGCTGTCTGTAACTAATCTGCTAGAACT 6162
Qy 1000 ACCCTACAGTACCGAGTATTGTATACAGATGAATAGATCGGTATCTGTGACATGTGG 1059
Db 6163 ATCCAGATGTACATCTTATCAGATTTACGAGAGTACGATCTTACTTGGATATGGTGG 6222
Qy 1060 ACGGCGCATCGTGTCTGTAGTATACAGCCACTTTTGTCCGGCTAACTGAGAACTACC 1119
Db 6223 ACGGCGAGTGGCGCTGCTGATATCTGCAACTTCTGCCCCGCTAAGCTAGAAATTTACC 6282
Qy 1120 CAAGAGCATATGCTATTTTTCAGCCAGATTAAGATCAGCGCTCCATCGCTATPACGA 1179
Db 6283 CGAAAAAATCATGATAGAGGCCCGCAATATCCGAGTGGCTGTATCATCAGAGTACAGA 6342
Qy 1180 ATACATTCACAAATGTATTTGGCTGAGCTACTAAAGGAATTCGAACGTTTCCCAAAATGC 1239
Db 6343 ACACGCTACAAATGTCTCATTTGCCGCACTTAAAGAAATTCGAACGTCACGACGATGC 6402
Qy 1240 GAGATTTACTGTCTTGAATTCGCGGCAATTTAATGTTGATGTTTCAAGAAATACGAT 1299
Db 6403 GTGAATCTCCAACTAGGACTCAGGACATTCATGATGGAATGCTTTGAAAATTTGAT 6462
Qy 1300 GCAATGATGATCTGGGATACCTTTCCGATTAACCTTATTCGCTTAACTACAGAAACG 1359
Db 6463 GTAATGACGAGTATTTGGAGGAGTTCGCTCGGAGCCCAATTAGATTAACACTGAGTTTG 6522
Qy 1360 TTACGCAATATGTGACAAAGCTGAAGGGCCGAAGCCAGCAGCATTTGTTGGAAATCTC 1419
Db 6523 TCACCGCATATGTAGCTAGACTGAAGGCGCTTAAAGCGCGCGCACTATTTTGAAGAGCT 6582
Qy 1420 ATAAATCTAAAACCGTTGACAGAGATACCAATGATCAATTCGTTCATGATCTTAAAGAG 1479
Db 6583 ATTAATTTGTCCCATTTGCAAGAGAGTCCCTATGATGATGATTCATGACATGAAAAGAG 6642
Qy 1480 ATGTCAAGTTTACTCCCGGACGAAACATPACAGAGAGCGGCTTAAAGGTGCGATTATTC 1539
Db 6643 ACGTCAAGTTTACACAGGACGAAACACACAGAAAGAAAGCCGAAGTACAGTATAC 6702
Qy 1540 AGGCGCAGATCCCTTGTCTACCGGCTTACCTTTGGGGATCCATGGGAATTAAGTCGTA 1599
Db 6703 AAGCGCAGAACCCCTGGGAGATGCTTACTTATGCGGATTCACCGGGAATTAAGTCGTA 6762
Qy 1600 GACTGAATCGGTGCTTTCGCAAAATATCATCTCTTCGACATGTCAGCGGAGAAAT 1659
Db 6763 GGCCTTAAGCGCGTCTTGTCTCAAACTTACACGCTTTTGAACATGTCGCGGAGGAT 6822
Qy 1660 TTGATGCAATTAATGCTGAACATTTTCACACGCGGACCCAGTATTTGAAACGACATCG 1719
Db 6823 TTGATGCAATCATAGCAAACTTCAAGCAAGCGGACCCGGTACTGAGAGACGATATCG 6882

Qy 1720 CGTCGTTGATTAAGCGAAGACGACGCTATCGCAATTTGCGGCTTGTATGATCCTTGAG 1779
Db 6883 CATCATTTGACAAAAGCCAAAGACGACGCTATGCGCTTAAACGGTCTGATGATCTTGAG 6942
Qy 1780 ACTTAGGTGACACCAACCGCTCTTGAATTTGATAGAGCGCGCTTCCGCAATATCAAT 1839
Db 6943 ACTGGGTGTGATCAACCACTACTGACTGATGATGAGTGGCGCTTTGGAGAAATATCAT 7002
Qy 1840 CTGTGACCTTACCTTACAGAAACGAGTTTAAATTTGGTCCATGATGTAATTCGGTATGT 1899
Db 7003 CCACCATCTTACCTTACGAGTACTGTTTAAATTCGGGCGATGATGAATTCGGAATGT 7062
Qy 1900 TCTTAACGCTGTTTGTCAACACACTAGTCAATATCATGATTTGCTGACAGACTACGTT 1959
Db 7063 TCTTCACCTTTTGTCAACACAGTGTGAATGTCTTATGCGCAGACGATGACTAGAG 7122
Qy 1960 AACGTTAACACGTCAGCGTGCAGCGGCTCTATTCGCGACGATTAACATGATGATGTT 2019
Db 7123 AGCGGCTTAAACGTCAGATGTGACGCTTCAATTTGGCGAGCAACATCATATGAGG 7182
Qy 2020 TCGTCTCGACACCTTGTATGCGGAGATGCGCACTTGTGCTGAACATGGAATTAATA 2079
Db 7183 TAGATCTGACAAAGAAATGCTGAGAGTGCAGCCTGCTCAACATGAGGTTAAGA 7242
Qy 2080 TTATGATGCAATTTATGTTTCAAGACCTTACTTGTGGGGGATTTATCCTGTGG 2139
Db 7243 TCATGACGAGTCACTCGGTAGAGACCACTTACTTCTGGCGGATTTATCTTGACAG 7302
Qy 2140 ACCAGATTAACAGGACCACTGACAGAGTGCAGACCTCTTAAAGAGCTTTTAAAGCTTG 2199
Db 7303 ATTCGTTACTTCCACAGCGTGGCGCGTGGGAGCCCTGAAAGAGCTTTAAAGTTGG 7362
Qy 2200 GAAACCACTTGGCAGTGCATGATACCAAGCTGCGACCGCGCGGCACTGCAATG 2259
Db 7363 GTAAACCGCTCCAGCCGACGACGAGAAAGCAAGACAGAAAGCCCTCTGCTAGATG 7422
Qy 2260 AAGCAATGCAATGGAACGAATTTGGAATTTAGGAGAGTGTGAAGGCCGAGAAATCCA 2319
Db 7423 AAACAAAGCGGTGTTTGAATGATGATTAACAGGCACTTTAGCACTGACGAGCCGAGAC 7482
Qy 2320 GATACGATATCATCTGACAGGCTGATCATCAGCTCTCTGTCCAGCTTAAAGCCGAAACG 2379
Db 7483 GGTATGAGTATGACAAATTTATACCTGTCTTACAGGATTTAGAACTTTTCCGAGAC 7542
Qy 2380 TTAAAGACTTAAGAGCATTAAGAGGAGCCCAATCATCTTACGCTGACCTTAAATAG 2439
Db 7543 AAAGGCAATTCGAAGCCATCAGAGGGGAAATTAAGCATCTTACGCTGCTCTTAAATAG 7602
Qy 2440 TGAGGTGATGACACGACCTTAC-----CCACCGGACGAATGTTTCCATACC 2486
Db 7603 CAGCATATATATTTTATCTGATTAATTAATTAACACACACCACTGAATGAGATTC 7662
Qy 2487 CTCAGCTGAACCTTTTCAACAGTTTTAACTTAACTTAACTTAACTTAACTTAACTTAA 2546
Db 7663 TTAACTATGCTGGCGCGCGCGCTTCCCGGCGCCCACTGCGCATGTGAGGCGCGGAGAA 7722
Qy 2547 CTCCTAGCGCGCTGAGCGCTTGGCGCTTGGCGCGCTGCTGCTCAATTCGAATCTTA 2606
Db 7723 GAGGCGAGCGCGCGCGCTGCTGCGCGCAACGCGGCTGCTTCTCAATTCAGCACTGA 7782
Qy 2607 GAGGTGATATGATCACTTGAATTTCAACAAACATCACTTAACTCCGCGCAGTCTCA 2665
Db 7783 CCAAGCGCTGATGCTTACTGATCATTTGACAGGCACTGACCTTAACTCCCACTCAC 7842
Qy 2666 -----CCGCAAGAGAGAGAGTGTCTTAAAGCCAAACCTTACTAGGCTTAAAGAG 2720
Db 7843 GCCCGCAACCGCGCGAG 7902
Qy 2721 AGAAGCAGCAAGCAAG 2777
Db 7903 AAAAGCAG 7962

Db 10123 AAGGGGAGGGGTATGCCCCGCTCAATTGGAGATCATGTCTATGCTCGAGGTTTTGC 10182
Qy 4983 CTTCAACTACAGAGAGTAGCTGACCTGCAATTCCACAGATCATTCCTTACACAG 5042
Db 10183 CTTCCACCAACAGAGAGTAGCTGACCTGCAATTCCACAGATCATTCCTTACACAG 10242
Qy 5043 TTAAATGCTGGGGTCCCTCGATGCAAGGATCTCAAGGCGGATTAACATGCCCCG 5102
Db 10243 TCAATGCTGGGGTCCCTCGATGCAAGGATCTCAAGGCGGATTAACATGCCCCG 10302
Qy 5103 TTTTGGGGGTGTGATCCCTTTCATGTGGGAGGGGCAATGCTTCTGTGACAGTGA 5162
Db 10303 TCTTGGAGGGGTGTGATCCCTTTCATGTGGGAGGGGCAATGCTTCTGTGACAGTGA 10362
Qy 5163 ACACACAATGATGAGGAGCTAGCTGAGTTGCTCCAGACTGCACTATAGATCAGGAG 5222
Db 10363 ACAGCCAGATGATGAGGAGCTAGCTGAGTTGCTCCAGACTGCACTATAGATCAGGAG 10422
Qy 5223 TCGCACTAAAGTTTACACAGCTGCTGTAAAGTGGGCTGATATGATATCGGCAACA 5282
Db 10423 AGGCGATTAAGGTGACACTGCGCGATGAAGTAGAGCTGGTATATGATCGGGAACA 10482
Qy 5283 CCACCGCGCACCCTGATACGTTTGTCAATGGCGTACGCCAGTTCTTCACGGGACTGA 5342
Db 10483 CTACCAAGTTTCTTATGATGTATGCTGAAAGGAGTACACAGGAACTGTAAGACTTGA 10542
Qy 5343 AGGTATAGAGGGGCGATATCAGCGGCTTTTACCCCTTGAACCATTAAGTCCGATCA 5402
Db 10543 AAGTATATGCTGAGCAATTTCAAGATGTTTACGCCATGATATTAAGTCTGTTATCC 10602
Qy 5403 GAAAGGGGCTTTTACAACTACGACTTCCCTGATGATGAGCTATGAACAGAGGCGT 5462
Db 10603 ATCGGGGCTGGTGTGACAACTATGACTTCCCGAATATGAGCGATGAACAGAGGCGT 10662
Qy 5463 TCGCGATATTAAGCATCTCGCTGTATGCTACAGACATAGTACCGGCACTGACATAC 5522
Db 10663 TTGAGACATTAAGCATCTCGCTGTATGCTACAGACATAGTACCGGCACTGACATAC 10722
Qy 5523 GGTGCTGAAGGCTTTGTGCAAGAACATCCAGTCCCTACACCCAGACATATCAGGCT 5582
Db 10723 GGTATCTCAAGGCTTTGTGCAAGAACATCCAGTCCCTACACCCAGACATATCAGGCT 10782
Qy 5583 ATGAATGTGAGAGAACACTCAGGACGAGCCCTGCAAGAAACAGACATTTGGATGTA 5642
Db 10783 TTGAGATGTGAGAAACAACACTCAGGACGAGCCCTGCAAGAAACAGACATTTGGATGTA 10842
Qy 5643 AAATTGAATGAGGCTTGTGAGAGGCTTAACTGTGCTTACGGGCAATCCCTATCTGA 5702
Db 10843 AGATTGCAATTAATCCGCTCCGAGGCTGAGCTGTTCAATCGGGAACATTCCTATCTGA 10902
Qy 5703 TTGACATCCCTGATGAGGCTTTGTGATCATGATCATCAACAACTTTAGAGTTA 5762
Db 10903 TTGACATCCCTGATGAGGCTTTGTGATCATGATCATCAACAACTTTAGAGTTA 10962
Qy 5763 GCTGACAGTAGACAGACTGATTTATTTGAGACTTTGTGTTCTCTTAACATTACAGT 5822
Db 10963 AATGTGAAGTAGAGAGTGACTTATTTAGAGACTTTGGCGGAGTGGCCACTGCAAGT 11022
Qy 5823 ACAAGCTGACAGGAGGAGCATTTGTCAGTTCACTCCACTCCACGACAGCTGTTTTGA 5882
Db 11023 ATGTATCCGACCGGAGAGGTCAATGCCCCGTAATTCGATTTGAGACACAACTCTCC 11082
Qy 5883 AGGAAGGACCAACATGTGACTGCCGTAAGCAGATTAACACTATTTTAGCATGTA 5942
Db 11083 AAGAGTGCACAGTACATGCTCTGAGAAAGAGCGGTACAGTACACTTTAGCACCGGGA 11142
Qy 5943 GCCCAAGCAAAATTTTATGTTTCTATGCTATGCGGCAAGAGTCCACTGCAATGCTGAT 6002
Db 11143 GTCCACAGGCGAATTTATGCTATGCTGTGTGGAGAAAGACAAATGCAATGCAAGAT 11202
Qy 6003 GTAAACCAACCGGCGACCACTAATTTGGAAGCAATTAAGTGCACCAAGATTTCCAG 6062
Db 11203 GTAAACCAACCGGCGACCACTAATTTGGAAGCAATTAAGTGCACCAAGATTTCCAG 11262

Qy 6063 CGGCAATTTCCAAAACATCTTGAAGTGGCTGTGCACTGTTTGGGGAGCATATCC 6122
Db 11263 CCGCATCTCAAAAACATCATGAGTGTGGCTTTGGCTTTTGGGGCGCTCTGTGC 11322
Qy 6123 TCATTTGTAGAGCTTATAGTGTGTGCTGACCTTATGCTTATTAACACAGATGAT 6182
Db 11323 TATTAATTAATGAGCTTATGATTTTGTCTGACAGATGATGCTGATAGCACAGAGAT 11382
Qy 6183 GACTGAGCGGGAACACTGA-CATAGCGGTAAACTCGATGTTCTTCCGAGGAGCGTGT 6241
Db 11383 GACCGTACGCCCCCAATGATTCGACAGCAAACTCGATGTTCTTCCGAGGAGCTGATGT 11442
Qy 6242 GCATATGCGCACGCGCGCT 6261
Db 11443 GCATATGCTATGAGCTGTGT 11462

Search completed: November 16, 2003, 05:39:13
Job time : 22826 secs

Result No.	Score	Query Match	Length	ID	Description
C 1	49.8	1.0	885	BX425603	BX45603
C 2	43.7	0.9	1201	BX381961	BX381961
C 3	43.8	0.9	1098	BX436460	BX436460
C 4	42	0.8	1201	BX381961	BX381961

C	5	41.2	0.8	646	14	CD393733	CD393733 CM CK1332
C	6	40.8	0.8	398	12	BG984025	BG984025 IL5 -CN006
C	7	40.8	0.8	971	13	BU173419	BU173419 AGENCOURT
C	8	40.8	0.8	1101	29	CNS01757	AL108460 Drosophila
C	9	40.6	0.8	380	12	BJ198034	BJ198034 BU198034
C	10	40.6	0.8	455	12	BJ198451	BJ198451 BU198451
C	11	40.6	0.8	453	12	BJ198768	BJ198768 BU198768
C	12	40.6	0.8	455	12	BJ203653	BJ203653 BJ203653
C	13	40.4	0.8	715	12	BJ195766	BJ195766 60274145
C	14	40.4	0.8	1201	13	BX376057	BX376057 BX376057
C	15	40.2	0.8	634	10	BE017217	BE017217 K7HB02.y
C	16	39.8	0.8	413	14	CB84158	CB84158 TGE872yh5
C	17	39.8	0.8	916	10	BG541362	BG541362 602568657
C	18	39.6	0.8	343	9	AMS01797	AMS01797 UT-HF-BRO
C	19	39.6	0.8	402	9	AMS01918	AMS01918 UT-HF-BRO
C	20	39.6	0.8	449	9	AM496678	AM496678 UT-HF-BRO
C	21	39.6	0.8	480	9	AM496647	AM496647 UT-HF-BRO
C	22	39.6	0.8	491	9	AM502451	AM502451 UT-HF-BRO
C	23	39.6	0.8	859	14	CA971504	CA971504 AGENCOURT
C	24	39.6	0.8	1011	10	BE563150	BE563150 60135780
C	25	39.4	0.8	504	13	BU266381	BU266381 sa890a08.
C	26	39.4	0.8	597	12	BM308456	BM308456 sak46c12.
C	27	39.4	0.8	619	13	BU927744	BU927744 sac13b07.
C	28	39.2	0.8	999	12	BM477744	BM477744 AGENCOURT
C	29	39	0.8	243	14	CB931641	CB931641 OSTF156A8
C	30	39	0.8	560	28	AZ738168	AZ738168 RPCT-24-1
C	31	39	0.8	599	13	BQ873224	BQ873224 OC117P07.
C	32	39	0.8	626	13	BQ849552	.BQ849552 QGB10C11.
C	33	39	0.8	997	29	CNS0057E	AL060757 Drosophila
C	34	39	0.8	1240	10	BG739729	BG739729 602630430
C	35	38.8	0.8	342	12	BI177442	BI177442 ESTS18387
C	36	38.8	0.8	650	28	BH900647	BH900647 Octb00189
C	37	38.8	0.8	687	12	BM406647	BM406647 ESTS80974
C	38	38.8	0.8	705	12	BI432608	BI432608 ESTS53639
C	39	38.8	0.8	718	10	BG598366	BG598366 ESTS03266
C	40	38.8	0.8	1101	29	CNS0176S	AL108022 Drosophila
C	41	38.8	0.8	1182	9	AL513775	AL513775 AL513775
C	42	38.4	0.8	413	13	BQ281915	BQ281915 WHE3026.H
C	43	38.4	0.8	1003	29	CNS04HPD	AL280374 Tetradon
C	44	38.4	0.8	1076	29	CNS05HKN	AL381810 Tetradon
C	45	38.4	0.8	1163	29	CC290181	CC290181 CH261-100

ALIGNMENTS

[illegible]

BASE COUNT	206 a	134 c	134 g	172 t	host cells (Stratagene)."
Query Match	0.8%	Score 41.2;	DB 14;	Length 646;	
Best local Similarity	48.9%	Pred. No. 3.4;			
Matches 139;	Conservative	0;	Mismatches 143;	Indels 2;	Gaps 1
Qy	2331	AGTGCCTCTTAATGGGGTTAAAGCACC	CGTTAAACATCTGTACATTTGATGAGCATTTGGC	2399	
Db	599	AGTCTCCCAATCAATAGAGGGGTTTAA	GATCATAGTGTGTGTTTGGCAACATGGCTTAA	540	
Qy	2391	CTGCCATGCAAGGAGCGCTGTGCACT	GATTTGCCATGTCTCAACCTTAAGAAAGTGTAT	2455	
Db	539	GTTCTTTGCTGTATGATCTTTGGCAGC	CTTGTGCATTTTCCATCTTCAAAACAGTGTAT	480	
Qy	2451	GTGGGGGACCCAAAACAATGCGGCTT	CTTTAAATATATGTCCTGGAAGTACATTTTAA	2511	
Db	479	GGCTCTTAATGAACAAAGAGGCCACT	GTATGACAACAGCAATATGTGAAGTGGAAG	420	
Qy	2511	CCATGACATATGCACTGAAGTATCCAT	TAAAGCATCTCTAGAGGTGACACAGACTGT	2577	
Db	419	TGTCAAGATTAACCGTAAATGCCAT	CAACAATG--CTCGAGGATGTATGACGACCA	362	
Qy	2571	AACCGCATGTCTTCCACGCTCTTCT	CTACGACACACGCAATGAGA 2614		
Db	361	ATACCAACAGCCCTGCAATGTTTCT	GTCTGCAAAAGTGTGACAGA 318		
RESULT 6	BSG984025/c	398 bp	mRNA	linear	EST 12-JUN-2001
LOCUS	IL5-CN0065-190301-403-d07	CN0065	Homo sapiens	CDNA,	mRNA sequence.
DEFINITION	BSG984025.				
ACCESSION	BSG984025.1	GI:14386760			
VERSION	EST.				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 398)				
AUTHORS	Dias Neto,E., Garcia Correa,R., Veijovaki-Almeida,S., Britones,M.R., Nagai,M.A., da Silva,M.J., Zazo,M.A., Bordin,S., Costa,F.P., Grindem,C.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.				
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)				
MEDLINE	20202863				
PUBMED	10737800				
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&ct=IL5-CN0065-190301-403-d07&et3=2001-03-19&ct4=1) Seq primer: puc 18 forward High quality sequence stop: 370. Location/Qualifiers 1. .398 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /dev_stage="Adult" /clone_11b="CN0065"				
FEATURES	source				

QY 1474 GATGTGAGCATCTTCGTCCTTACAGCAAGAAGCTGAGAGTGGCTGCACGGAGAG 1533

1474 GAGTGGAGCATCTGCGTGGCTTACAGCAAGAACTGAAGAAGTGGCTGCACGGCAAGAG 1533

Db	103	GAGGAGGAGGAGGAGGAATTTCAGACCAGCTTTGAAGAAAGAAAGAAAGAAAGA	162
QY	1534	ATCAGAGAAGCCCTCCACCCCTGCTCCCTGAAATAGAAAAAGAACCGTAGAGCCAGA	1593
Db	163	GAAAGCTGTATGTCGAAGCAAGACTCCAGACGAAAGAAAGAAAGAAAGCTGTATGTCGAACCA	222
QY	1594	GTAGACCTCATTTATGCAAGAGCGAGAGAGGTAACTCTGAGACACACACAGAGACATC	1653
Db	223	GAAAGTCGACGAAGAAAGAAAGAAAGAACAGCAGGAGGAGAGACCGCATAGAAAGAAAGA	282
QY	1654	AGGCTGACAAAG	1664
Db	283	AGAAAGAGAAAG	293

RESULT 11

LOCUS	DEFINITION	Accession
Bj198768	455 bp mRNA linear	EST 24-JAN-2002
Bj198768	normalized full length cDNA library, chironomata, caulonemata and rhizoid-like protoneurata Physcomitrella patens subsp. patens cDNA clone pphm22c20 5', mRNA sequence.	
Bj198768		

ACCESSION	BJ198768
VERSION	BJ198768.1
KEYWORDS	GI:18366689 EST.

SOURCE ORGANISM	physcomitrella patens subsp. patens	physcomitrella patens subsp. patens
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

REFERENCE

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella 1 (bases 1 to 455)

AUTHORS
Fujita, T., Shin-¹, Seki, M., Kamiya, A., Uchiyama, J., Nishiyama, T., Carmaci, P., Hayashizaki, Y., Shinozaki, K., Kohara, Y. and Hasebe M.

TITLE	Comparison of the moss <i>Physcomitrella patens</i> genome with flowering
JOURNAL	plants genome
COMMENT	unpublished
	Contact: Tadao Shin-i

Center for Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855

Email: tschini@genes.nig.ac.jp
A backbone of the vector is pBluescript II, that was in vivo excised from a modified lps phage vector (Wo bi Tec, Germany). XhoI digested 5' end of cDNA is ligated to SalI site of the vector, and the BamHI digested 3' end including polyA tail is ligated to BamHI site of the vector. cDNA insert could be amplified with conventional T7 and T3 primers. This normalized full-length cDNA library was generated basically according to the method described in Genome Research 10, 1617-1630 (2000). Carninci, P. et al. Protonemata were blended by the POLYTRON and then cultivated on the BCD medium containing 1mM NAA (naphthalene acetic acid) for 8 to 11 days under the continuous light.

FEATURES
source

Source

```

/organism="Physcomitrella patens subsp. patens"
/mol_type="mrna"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="rphn3c20"
/tissue_type="mixture of chloronemata, caulonemata and
rhizoid-like protonemata"
/clone_lib="normalized full length cDNA library,
chloronemata, caulonemata and rhizoid-like protonemata"
BASE COUNT      138 a      134 g
ORIGIN           111 c      72 t

```

BASE COUNT
ORIGIN

Query Match	0.88;	Score 40.6;	DB 12;	Length 455;
-------------	-------	-------------	--------	-------------

Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

1474 GATGTGAGCATCTGCGGTACAGCAAGAGCTGAAGAGTGGCTGCAGCGGAAGAG 1533

11

Db 103 GAGGAGGAGGAGCAGAGATTACGACCAGCAGTTGAGAGAGAGAGAGAGAGAGAA 162

1534 ATCAGAGAAGCCCTGCTCCCTGAATAGAAAAGAGACCGTAGAGGCAGAA 1593

Db 163 GAAGCTGTAGTCGAAGCAGAGTCGACGAGAAGAGAGAGAGCTGTAGTCGAAGCA 222

1594 GIAGACCTCATTAAGCAAGAGCAGGACGAGTACCGTGGAGACACCGAGGACACATC 1655

[illegible][illegible]

Db 283 AGAAGGAGAAG 293

RESULT 12

LOCUS B2J03653 456 bp. mRNA linear EST 25-JAN-2002
DEFINITION B2J03653 normalized full length cDNA library, chloronemata,
calionemata and rhizoid-like procionemata Physcomitrella patens
subsp. patens cDNA clone pphm49314 5', mRNA sequence.

ACCESSION	BJ203653
VERSION	BJ203653.1
KEYWORDS	EST.

SOURCE ORGANISM	Physcomitrella patens subsp. patens	Physcomitrella patens subsp. patens
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

REFERENCE
1 (bases 1 to 456)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.

AUTHORS
Fujita, T., Shin-i, T., Seki, M., Kamiya, A., Uchiyama, I., Nishiyama, T.,
Carinci, P., Hayashizaki, Y., Shinzaki, K., Kohara, Y. and Hasebe
M.

TITLE	Comparison of the moss <i>Physcomitrella patens</i> genome with flowering
JOURNAL	plants genome
COMMENT	unpublished
	Contact: Tadasu Shin-i

Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855

Email: tehini@genes.nig.ac.jp
A backbone of the vector is pBluescript II, that was in vivo excised from a modified lPS phage vector (Mo bi Tec, Germany). XbaI digested-5' end of cDNA is ligated to SalI site of the vector, and the BamHI digested-3' end including poly-A tail is ligated to BamHI site of the vector. cDNA insert could be amplified with conventional T7 and T3 primers. This normalized full-length cDNA library was generated basically according to the method described in Genome Research 10, 1617-1630 (2000). Carrinoli, P. et al. Proteome data were blended by the POLYTRON, and then cultivated on the BCD medium containing 1 mM NMA (naphthalene acetic acid) for 8 to 11 days under the continuous light.

FEATURES

Source

```

/organism="Phrycomitrella patens subsp. patens"
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pphn49j14"
/tissue_type="mixture of chloronemata, caulonemata and rhizoid-like protonemata"
/clone_1fb="normalized full length cDNA library",
chloronemata, caulonemata and rhizoid-like protonemata"
BASE COUNT      138 a      134 g      73 t
ORIGIN

```

BASE C
ORIGIN

Query Match	0.88	Score 40.6	DB 12	Length 456
-------------	------	------------	-------	------------

Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

1474 GATGTGAGCATCTGCGTGGCTTACAGCAGAAGCTGAAGAAGTGGCTGCAGCGAAGAG 1533

100

Db	103	GAGAGGAGGAGGAGCAGAGATTTCAGACACAGAGCTGAAGAAAGAAAGAAAGAAAGAA	162
Qy	1354	ATCAGAGAACCCCTCCACCCCTGCTCCCTGAATATGAAAAAGAGACCGTAGAGCAGAA	1593
Db	163	GAAGCTGTAGTCGAAGCAGAAAGTCAGCAAGAAAGAAAGAAAGAAAGCTGTAGTCAGAACCA	222
Qy	1594	GTAGACCTCATTTATGACAGAGCAGAGCAGAGGTAGCGTAGAGACACACACAGAGCACATTC	1655
Db	223	GAAGTCGACGAAGAAAGAAAGAAAGAACAGCAGAGAGAGAGCAGCATTAAGAAAGAAAGAA	282
Qy	1654	AGGCTGACAAAG 1664	
Db	283	AGAAGGAGAAAG 293	
RESULT 13			
LOCUS	B1195766	715 bp	mRNA linear EST 10-JUL-2001
DEFINITION	602754145P1 NIH_MGC_19	Homo sapiens	CDNA clone IMAGE:4889938 5',
ACCESSION	B1195766		mRNA sequence.
VERSION	B1195766.1	GI:14650786	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 715)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: ATCC		
	CNA Library Preparation: Ling Hong/Rubin Laboratory		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Gene distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: image.lnl.gov		
	Plate: L16M1766 row: 1 column: 11		
	High quality sequence stop: 556.		
FEATURES			
source	location/Qualifiers		
	1..715		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone_image="4889938"		
	/tissue_type="neuroblastoma"		
	/lab_host="DH10B (phage-resistant)"		
	/clone_lib="NIH MGC 19"		
	/note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5'-adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."		
BASE COUNT	202 a 194 c 251 g 68 t		
ORIGIN			
Query Match	0.8%; Score 40.4; DB 12; Length 715;		
Best Local Similarity	50.5%; Pred. No. 6.5;		
Matches	98; Conservative 0; Mismatches 96; Indels 0; Gaps 0;		
Qy	1445	TCAAACCGCACCGGCTATTACATATGCGCATGTGAGACATCTCGTGCGTTACAGCAAG	1504
Db	223	TTTAAACAGCAGACAGCAGCTCAGACAGCAGCAGCAGACCCCGAGACACACATCAAA	282
Qy	1505	AAGCTGAAGAAGTGGCTGACCGAGCAAGAGATCAAGAAAGCCCTGCAACCTTGTCTCCTG	1566
Db	283	CACCTGTGCACCAACGGCAGCGGGCGCATTAAGAGAGCAAGAGAGAGAGCGCGCGCTG	342

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 22:31:43 ; Search time 12073 Seconds
(without alignments)
13055.125 Million cell updates/sec

Title: US-10-023-649-1_COPY_5000_11484

Perfect score: 6485

Sequence: 1 caccgcgcagcagcagcacc.....catttgccttaaatc 6485

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estnu:*
4: em_estmu:*
5: em_estcov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrt:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	58	0.9	985	13	EX436838
2	53.6	0.8	836	29	CNS04Q18
3	53.4	0.8	1101	29	CNS003DQ
4	52	0.8	1146	29	CNS021G2

5	50.8	0.8	1201	13	EX358198
6	50.4	0.8	997	29	CNS005TE
7	50.2	0.8	1101	29	CNS0182P
8	49.4	0.8	460	14	CB472367
9	49	0.8	844	29	AG042177
10	49	0.8	1101	29	CNS00E7Y
11	49	0.8	1165	13	EX338369
12	48.8	0.8	1198	13	EX395612
13	48.8	0.8	1201	13	EX385348
14	48.2	0.7	440	13	BU497603
15	48.2	0.7	566	12	BM275442
16	48.2	0.7	569	12	BM273626
17	48	0.7	887	13	EX415520
18	47.8	0.7	946	13	EX459436
19	47.6	0.7	993	13	BO729165
20	47.4	0.7	1065	9	AL515373
21	47.4	0.7	1101	29	CNS0021U
22	47.4	0.7	1225	29	CNS0161D
23	47.2	0.7	240	9	AU074526
24	47.2	0.7	944	29	CNS06WVA
25	47.2	0.7	994	13	EX414650
26	47.2	0.7	1005	13	EX414452
27	47.2	0.7	1201	9	AL548154
28	47	0.7	712	13	EX416727
29	47	0.7	1201	9	AL558113
30	46.8	0.7	905	29	CNS00KHX
31	46.8	0.7	913	28	AZ546811
32	46.8	0.7	975	13	EX419927
33	46.8	0.7	994	28	CNS015XG
34	46.6	0.7	589	28	AQ541309
35	46.6	0.7	794	28	BH498125
36	46.6	0.7	964	29	CNS017VX
37	46.6	0.7	1101	29	CNS0183Y
38	46.6	0.7	1161	13	EX332029
39	46.6	0.7	1201	13	EX396692
40	46.4	0.7	923	28	AZ538694
41	46.4	0.7	1101	29	CNS003DX
42	46.4	0.7	1201	13	EX336719
43	46.4	0.7	1201	29	CNS016AD
44	46.2	0.7	1101	29	CNS003DQ
45	46.2	0.7	1167	29	CNS07360

ALIGNMENTS

RESULT 1
EX436838
LOCUS
DEFINITION
BX436838 Homo sapiens THYMUS Homo sapiens CDNA clone CS0CAP003YH18
3-PRIME, mRNA sequence.
ACCESSION
EX436838
VERSION
BX436838.1 GI:30789527
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
Li W.B., Gruber C., Jessee J. and Polayes D.
Full-length cDNA libraries and normalization
Unpublished
CONTACT
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifeotech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0CAP003D09NP1.

FEATURES
source

location/Qualifiers
1..985
/organism="Homo sapiens"

```

/mot_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0CAP003YH18"
/cisue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

BASE COUNT      337 a      120 c      95 g      277 t      156 others
ORIGIN

Query Match      0.9%; Score 58; DB 13; Length 985;
Best Local Similarity 46.7%; Pred. No. 5.5;
Matches 84; Conservative 27; Mismatches 69; Indels 0; Gaps 0;

Qy 6303 TAAATGCTGTCAGTGCATTAATCGATATTCACATACATATTAACAACATATATCA 6362
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 746 WMAATTTTAAATAAATAATATATATATATATATATATATATATATATATATATATAT 805
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 6363 CTTTATGACACTGCTATGCGTCTTAATATACATACATACATATTTTACTTAAAAAC 6422
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 806 ATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 865
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 6423 TATACACATTAATTAATTTCTTATATATTTCTTTGTTTATTTGTTTAAAT 6482
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 866 ATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 925
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
CNS04Q18/c
LOCUS      CNS04Q18      836 bp      DNA      linear      GSS 01-SEP-2000
DEFINITION Tetradon nigroviridis genome survey sequence. PUC-ori end of clone
              188L13 of library G from Tetradon nigroviridis, genomic survey
              sequence.
ACCESSION   AL302021.1 GI:8179132
VERSION     GSS; genome survey sequence.
KEYWORDS    Tetradon nigroviridis
SOURCE      Tetradon nigroviridis
ORGANISM    Tetradon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1
AUTHORS     Roest Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fzames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
            Estimate of human gene number provided by genome-wide analysis
            using Tetradon nigroviridis DNA sequence
            Nat. Genet. 25 (2), 235-238 (2000)
JOURNAL     20296633
MEDLINE     10835645
PUBMED

REFERENCE   2
AUTHORS     Roest Crolius,H., Jallion,O., Dasilva,C., Ozouf-Costaz,C.,
            Fzames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
            Saurin,W., Bernot,A. and Weissenbach,J.
            Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetradon nigroviridis
            Genome Res. 10 (7), 939-949 (2000)
JOURNAL     20359837
MEDLINE     10899143
PUBMED

REFERENCE   3 (bases 1 to 836)
GENOSCOPE.
TITLE       Direct Submission
JOURNAL     Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            Web : www.genoscope.cns.fr)
            This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetradon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/tetradon.
COMMENT

```

```

Location/Qualifiers
1..836
/organism="Tetradon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="128L13"
/clone_lib="G"
/notes="Genoscope sequence ID : COB8128CF07SP1-end :
PUC-ori"

BASE COUNT      301 a      30 c      173 g      299 t      33 others
ORIGIN

Query Match      0.8%; Score 53.6; DB 29; Length 836;
Best Local Similarity 57.8%; Pred. No. 33;
Matches 89; Conservative 3; Mismatches 62; Indels 0; Gaps 0;

Qy 6331 ATATACACTATCATATTAAACAACATATATCTTATAGACTCATGCTCTCTA 6390
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 579 ATTTAATCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 520
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 6391 ATATACACTATCATATTATTTACTTAAACATATACATCTTATTAATTTCTTTAA 6450
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 519 ATACTAATTTTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 460
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 6451 TTTTCTTTGTTTATTTGTTTAAATTT 6484
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 459 TTATTTTATTTATTTATTTTTCATTTTTCATTTT 426
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 3
CNS003DQ
LOCUS      CNS003DQ      1101 bp      DNA      linear      GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR08109 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL064580.1 GI:4941932
VERSION     GSS.
KEYWORDS    Drosophila melanogaster (fruit fly)
SOURCE      Drosophila melanogaster
ORGANISM    Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 1101)
GENOSCOPE.
TITLE       Direct Submission
JOURNAL     Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            Web : www.genoscope.cns.fr)
            Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mamoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            pl and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
            Location/Qualifiers
            1..1101
            /organism="Drosophila melanogaster"
            /mol_type="genomic DNA"
            /db_xref="taxon:7227"
            /clone="BACR08109"
            /clone_lib="RPCI-98"
            /note="end : TET3"

BASE COUNT      291 a      51 c      117 g      404 t      238 others
COMMENT

```

BASE COUNT ORIGIN	a	c	g	t	others
	373	144	192	367	125

QY 2672 AAGAGAAGAGAGTCTCTTACGCCAAACCTACTCAGCTAAAGAAGACAGCAAA 273

Db 953 AAGAGAAGAAGAGAGAGAGARAAAARRRRRAGAGAGRGSGAGGGRGGRRAAA 894

BASE COUNT	274 a	268 c	128 g	73 t	358 others
ORIGIN					

[illegible]

Oy	6444	TTTAAATTTCTTTGTTTATTGTTTAAAA	6481
Dd	421	TTAAAATAAGTACTTTTTTAATTTTTTAAATAA	384
RESULT 10			
CNS00E7Y		1101 bp	DNA
LOCUS			
DEFINITION			
ACCESSION	CNS00E7Y		
VERSION	BACR28A16 of RPCL-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
KEYWORDS	A1069057		
SOURCE	AL069057.1 GI:4949017		
ORGANISM	GSS.		
AUTHORS	Drosophila melanogaster (fruit fly)		
TITLE	Drosophila melanogaster		
JOURNAL	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyridae; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 1101)		
COMMENT	Genoscope. Direct Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Osoegawa and Aaron Mammoss in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCL-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.		
FEATURES			
Source	Location/Qualifiers 1..1101 /organism="Drosophila melanogaster" /mol_type="Genomic DNA" /db_xref="taxon:7227" /clone="BACR28A16" /clone_id="RPCL-98" /note="end : TET3"		
BASE COUNT	248 a 175 c 173 g 341 t	164 others	
ORIGIN			
Query Match	0.8%; Score 49; DB 29; Length 1101;		
Best Local Similarity	38.1%; Fred. No. 1.9e+02;		
Matches	61; Conservative 44; Mismatches 55; Indels 0; Gaps 0;		
Oy	6318	TCACATTAATGTATATCACACTACATATTAACAACACATATATCTTTATGAGACGCA	6377
Dd	910	TCCMHAAHCHMHAHTHCHTMWMTATTTTMAATTTTWTMTTTTTTTTATTTTWA	969
Oy	6378	CTATGGCTCTGAATATACACTACACATATTTTACTTAAAAAACACTATACACTTTATA	6437
Dd	970	AMHTTHMMTMAAAATATTTTATTAHAAATMTMTTTTMMAMWMTAMWHHTCMVNY	1029
Oy	6438	AATCTTTATTAATTTTCTTTTGTTTATTTTGTTTT	6477
Dd	1030	CMTTTTTTTWTMTMTHTMTMTTTTTTTTTTTTTTTTTTTTTTTTTTT	1069
RESULT 11			
LOCUS	BXJ38369	1165 bp	mRNA
DEFINITION	BXJ38369 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA		
EST 02-MAY-2003			

ACCESSION	clone CSODI058YF14 5-PRIME, mRNA sequence.
VERSION	BX338369
KEYWORDS	BX338369.1 GI:30343682
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS	L.I.W.B., Gruber.C., Jessee.J. and Polyses.D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished
COMMENT	Contact : Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CSODI058DC070P1. Location/Qualifiers 1. 1165
FEATURES	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CSODI058YF14" /issue_type="PLACENTA COT 25-NORMALIZED" /clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED" /note="1st strand cDNA was primed with a NotI-cligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT	567 a 109 c 140 g 251 t 98 others
ORIGIN	
Query Match	0.8%; Score 49; DB 13; Length 1165;
Best Local Similarity	41.1%; Pred. No. 1.8e+02;
Matches	67; Conservative 33; Mismatches 63; Indels 0; Gaps 0;
Dy	6322 ATTATCGTATATCAGCTACTCATTTAACAACACTATATCATCTTTATGAGACTTCATAT 6381
Dd	:: ::
Dy	977 WMTWTWMAAAAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATA 918
Dd	:: ::
Dy	6382 GGCGTCTAATATACACTACACTATTCTTAATTAATAAACACTATACACACTTTAATAT 6441
Dd	917 WTTTATTATTATTTTAAATATWTTTATTATWAAAAATATATWTWMAAAAAWAATAATT 858
Oy	6442 CTTTATATAATTTCTTTGTTTTTATTTGTTTTTAAATTT 6484
Dd	:: ::
Oy	857 YTTTWAAATATWATWATAATATWTTTTTTTTTTTTTTTTTT 815
RESULT 12	
BX395612/c	
LOCUS	BX395612 1198 bp mRNA linear EST 13-MAY-2003
DEFINITION	BX395612 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
ACCESSION	Clone CSODI002YP22 3-PRIME, mRNA sequence.
VERSION	BX395612
KEYWORDS	BX395612.1 GI:30616740
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS	L.I.W.B., Gruber.C., Jessee.J. and Polyses.D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished
COMMENT	Contact : Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of

mRNA isolation system (Promega, WI) using streptavidin
MagneSphere particles. Directional cDNA libraries were
constructed by oligo d(T) priming of poly(A)+ RNA (5mg)
into EcoRI and XhoI sites of 1 ZapII vector using the Zap
cDNA synthesis kit (Stratagene, CA). The average size of the
cDNA inserts in the library was between 1.0 and 1.5kb.
Clones were mass excised using the ExAssist helper phage
(Stratagene), the phagemids were precipitated with PEG
8000 and extracted with phenol/chloroform. Phagemid DNA
was electroporated into DH10B cells."

BASE COUNT 195 a 22 c 14 g 208 t 1 others
ORIGIN

Query Match 0.7%; Score 48.2; DB 13; Length 440;
Best Local Similarity 55.8%; Pred. No. 3.3e+02;
Matches 92; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 6320 ACATTAACTGATATCACTACATATTAAACACATATACCTTTATGAGCTCACT 6379
DB 215 AT 274
QY 6380 ATGGCTCTATATATACATACATATTTTACTTAAACACATATACACTTATATA 6439
DB 275 ATTAAT 334
QY 6440 TTTCTTATATATATTTCTTTGTTTATATTTGTTTAAATTT 6484
DB 335 ATATATATATATTTCTTATATATATATATATATATATATATTTT 379

RESULT 15 566 bp mRNA linear EST 20-DEC-2001
BM275442
LOCUS
DEFINITION BM275442.1 GI:17968772
P1EST00a68h12.y1 Plasmodium falciparum 3D7 gametocyte cDNA library
Plasmodium falciparum 3D7 cDNA 5' similar to SW:UBC1.DROME P25867
UBIQUITIN-CONJUGATING ENZYME E2-17 KD; mRNA sequence.

ACCESSION BM275442
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
Tang, K., Cole, R., Chakrabarti, D., Haywood, R., Clifton, S., Page, D.,
Matta, M., Hillier, L., Martin, J., Wylie, T., Dante, M., Theising, B.,
Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Jentes, E., Ronko, I.,
Tsagarashvili, R., Belaygorod, L., Franklin, C., Carr, L., Grow, A.,
Maguire, L., Richey, J., Wadkins, J., Kennedy, S., Levinso, D.,
Waterston, R., Wilson, R. and Sibley, D.
Washu Plasmodium EST Project
Unpublished

TITLE
JOURNAL
COMMENT
Contact: L. David Sibley
Washu Plasmodium EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314.286.1800
Fax: 314.286.1810
Email: est@wustl.edu

Library was constructed by R. Haywood. DNA sequencing by:
Washington University Genome Sequencing Center for information on
obtaining a clone please contact: L. David Sibley
(sibley@jorcm.wustl.edu), Washington University
Seq primer: -40UP from Gibco
High quality sequence stop: 423.

FEATURES

source
1. .566
Location/Qualifiers
/organism="Plasmodium falciparum 3D7"
/mol_type="mRNA"
/db_xref="taxon:36329"
/dev_stage="gametocyte (stage III-V)"
/lab_host="DH10B (GeneHog, Invitrogen, Inc.)"
/clone_lib="Plasmodium falciparum 3D7 gametocyte cDNA
library"

/note="Vector: pBluescript SK plus; Site 1: EcoRI; Site 2:
XhoI; The library was constructed by R. Haywood. cDNAs were
synthesized from gametocyte poly(A)+ RNA by oligo d(T)
priming, size-selected and directionally cloned into the
EcoRI (5' end) to XhoI (3' end) sites of the Uni-ZAP XR
lambda vector (Stratagene). The primary library was mass
excised as phagemid using the ExAssist helper phage
(Stratagene). Clones were mass excised using the ExAssist
helper phage (Stratagene), the phagemids were precipitated
with PEG 8000 and extracted with phenol/chloroform.
Phagemid DNA was electroporated into DH10B cells. Clone
Availability: David Sibley, Washington University."

BASE COUNT 256 a 63 c 194 t
ORIGIN

Query Match 0.7%; Score 48.2; DB 12; Length 566;
Best Local Similarity 57.7%; Pred. No. 3.1e+02;
Matches 86; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 6336 ACACTACATATTAACAACATATATACCTTTATGAGCTCACTATGGCTCAATATA 6395
DB 1 AGAATA 60
QY 6396 CACTACATATATTTTACTTAAACACATATACACTTATATAATCTTTATATATTTT 6455
DB 61 TAATATATTTATATATATTTTATTAACAATATATATATATATATATATATA 120
QY 6456 CTTTGTATATATTTGTTTAAATTT 6484
DB 121 TATATATATCTCTTTTTCACACTT 149

Search completed: November 16, 2003, 08:59:44
Job time : 12079 secs

Db 61 CCGATGCAAGTGTATACAGGACGTTTCCAACTTGTAGTGAAGCAAGGAGTTC 120
Qy 121 ACTGACAAATGACATGCCAATGCGAGAGCGTTTTCGATGTGGCAACAAGCTCATTTAG 180
Db 121 ACTGACAAATGACATGCCAATGCGAGAGCGTTTTCGATGTGGCAACAAGCTCATTTAG 180
Qy 181 AGCGAAGTCGACCGGAGCAAGTTATCTTGGACATGTGAAGTGGCCCGTCAGACATGCA 240
Db 181 AGCGAAGTCGACCGGAGCAAGTTATCTTGGACATGTGAAGTGGCCCGTCAGACATGCA 240
Qy 241 CATTCCAATCACCGCTATCATTTGATCTGCCCTATGATTAAGCGCTGAAGACCCGGACGA 300
Db 241 CATTCCAATCACCGCTATCATTTGATCTGCCCTATGATTAAGCGCTGAAGACCCGGACGA 300
Qy 301 CTACACGCGTATGCGAAGAAAGCTTAAGAAAGTGCATTACCGACAGACATAGCCTCT 360
Db 301 CTACACGCGTATGCGAAGAAAGCTTAAGAAAGTGCATTACCGACAGACATAGCCTCT 360
Qy 361 AAGGCGGAGACCTGCTGGAGTCAATGTCACACAGAGGAGACTCCATCTCTGTGT 420
Db 361 AAGGCGGAGACCTGCTGGAGTCAATGTCACACAGAGGAGACTCCATCTCTGTGT 420
Qy 421 ATGCAACAGACGCGCAGCTGATGATCTTGGAGTGTAGCAGTATACCAAGATGTGTAC 480
Db 421 ATGCAACAGACGCGCAGCTGATGATCTTGGAGTGTAGCAGTATACCAAGATGTGTAC 480
Qy 481 GCAATCCATGCAACCGACATCAATCTACACAGGCGCTTAAAGAGTTAGCAATTTAC 540
Db 481 GCAATCCATGCAACCGACATCAATCTACACAGGCGCTTAAAGAGTTAGCAATTTAC 540
Qy 541 TGGATAGGCTTGAACAGACCCCTTTATGTACAAAACATGGAGGTTCTCACTCACT 600
Db 541 TGGATAGGCTTGAACAGACCCCTTTATGTACAAAACATGGAGGTTCTCACTCACT 600
Qy 601 TACAACAGAACTGGGCTGACGAGAGAGTATGTGAAGCAGCTAACATTGGCCTGCTAAC 660
Db 601 TACAACAGAACTGGGCTGACGAGAGAGTATGTGAAGCAGCTAACATTGGCCTGCTAAC 660
Qy 661 TCAATCTTCAGAGAGAGCGCTTGGAAACTCTCAATCTTGAAGAAAGAGCTCCAA 720
Db 661 TCAATCTTCAGAGAGAGCGCTTGGAAACTCTCAATCTTGAAGAAAGAGCTCCAA 720
Qy 721 CCTACTAATAGATCATTTCTCGGTGTCTCAACATCTACAGAAAGATAGATCATG 780
Db 721 CCTACTAATAGATCATTTCTCGGTGTCTCAACATCTACAGAAAGATAGATCATG 780
Qy 781 TTACGTAGCTGGCATCTTCCAAAGCTGTTCACATTGAAAGAAAGTCACTTACAGAGT 840
Db 781 TTACGTAGCTGGCATCTTCCAAAGCTGTTCACATTGAAAGAAAGTCACTTACAGAGT 840
Qy 841 AGATGTGGAGCAATTTGCTGAGTGAAGGTTACGTCATCAAAAAGATACATGACGCCA 900
Db 841 AGATGTGGAGCAATTTGCTGAGTGAAGGTTACGTCATCAAAAAGATACATGACGCCA 900
Qy 901 GAGCTATACGCTTAAGTTGAAGCTTGGCGTCAACATGCAATCGAGAGGTTCTTGAAGT 960
Db 901 GAGCTATACGCTTAAGTTGAAGCTTGGCGTCAACATGCAATCGAGAGGTTCTTGAAGT 960
Qy 961 TGCAGAGTCACAGATACGCTCGCGGCGAGAGAGGTTCTTGTGCTGTGTACATGTA 1020
Db 961 TGCAGAGTCACAGATACGCTCGCGGCGAGAGAGGTTCTTGTGCTGTGTACATGTA 1020
Qy 1021 CCAGCCACACTTTGGCATCAAGTGAAGGAAATTCGCACTGACGTTAAGTGTGATGAC 1080
Db 1021 CCAGCCACACTTTGGCATCAAGTGAAGGAAATTCGCACTGACGTTAAGTGTGATGAC 1080
Qy 1081 GCAAGAAAACATATGCTTGGGCTCAACCAAGAGATGTGTCATAGTGAAGACGAAAGA 1140
Db 1081 GCAAGAAAACATATGCTTGGGCTCAACCAAGAGATGTGTCATAGTGAAGACGAAAGA 1140
Qy 1141 AATACTAAACAATGACAGAACTATCTATTACAGGTGTGCGCCAGGCGTTTTCAGGTG 1200
Db 1141 AATACTAAACAATGACAGAACTATCTATTACAGGTGTGCGCCAGGCGTTTTCAGGTG 1200

Db 1141 AATACTAAACAATGACAGAACTATCTATTACAGGTGTGCGCCAGGCGTTTTCAGGTG 1200
Qy 1201 GCGGTGAACATGCTGCCACTTGGAGACGAGAAAGAACTAAGGGTTCGGAGCGCACT 1260
Db 1201 GCGGTGAACATGCTGCCACTTGGAGACGAGAAAGAACTAAGGGTTCGGAGCGCACT 1260
Qy 1261 CTTACTATGGGCTGCTGCGGCTTTCAAGACCCAGAAATTCATCATCTACAAAG 1320
Db 1261 CTTACTATGGGCTGCTGCGGCTTTCAAGACCCAGAAATTCATCATCTACAAAG 1320
Qy 1321 CTTGTAACGCAACATTAAGAAAGTACCTGCGCTCTTGAATCTATTTGTGATTCAGC 1380
Db 1321 CTTGTAACGCAACATTAAGAAAGTACCTGCGCTCTTGAATCTATTTGTGATTCAGC 1380
Qy 1381 CTTTACAGCCACGGGCTCGATATGGGCTTCCCGGTAGGCTCAAGCTGCTGTGAACA 1440
Db 1381 CTTTACAGCCACGGGCTCGATATGGGCTTCCCGGTAGGCTCAAGCTGCTGTGAACA 1440
Qy 1441 ACTGTCAAAACCGGACCGGCTATTACAAATGAGCCGATGTGGAGCATGCGTGGCTACAG 1500
Db 1441 ACTGTCAAAACCGGACCGGCTATTACAAATGAGCCGATGTGGAGCATGCGTGGCTACAG 1500
Qy 1501 CAAGAACCTGAAGAGTGGCTGACGCGAAAGATCAGAGAGCCCTGCAACCTTGTCTC 1560
Db 1501 CAAGAACCTGAAGAGTGGCTGACGCGAAAGATCAGAGAGCCCTGCAACCTTGTCTC 1560
Qy 1561 CTTGAAATTAAGAAAGAGACCTGTAGAGGCGAAGTACCTCATTTATGCAAGGCGAGA 1620
Db 1561 CTTGAAATTAAGAAAGAGACCTGTAGAGGCGAAGTACCTCATTTATGCAAGGCGAGA 1620
Qy 1621 GCAAGTACGCTGAGACACCAAGGACATCAGGCTGACCAAGTTCCAGGCGAAGAG 1680
Db 1621 GCAAGTACGCTGAGACACCAAGGACATCAGGCTGACCAAGTTCCAGGCGAAGAG 1680
Qy 1681 AAGATGGGCTTACGCTATATCTTTCACCCAGGCGGTATGAAATATGAAAACTGGCG 1740
Db 1681 AAGATGGGCTTACGCTATATCTTTCACCCAGGCGGTATGAAATATGAAAACTGGCG 1740
Qy 1741 TGTATCCACCATTTGGGGGAAACAAAGTACTGGTATATGCTCAACAAAGGTAGGGCGAGGA 1800
Db 1741 TGTATCCACCATTTGGGGGAAACAAAGTACTGGTATATGCTCAACAAAGGTAGGGCGAGGA 1800
Qy 1801 TACAAGTGCAGCATCCACAGGTAAGTCAATGTATCAGAAAGGAGCGGCTCCCTGT 1860
Db 1801 TACAAGTGCAGCATCCACAGGTAAGTCAATGTATACCAAGGAGCGGCTCCCTGT 1860
Qy 1861 CAAGACTTCAAGGCAATTAAGTGAAGGCGTACGATCGTTTCAACGAGAGGAGTTGCTGA 1920
Db 1861 CAAGACTTCAAGGCAATTAAGTGAAGGCGTACGATCGTTTCAACGAGAGGAGTTGCTGA 1920
Qy 1921 AACAGATACCTGCAACATCGCAATCAACGAGAGGCGCTTAAACATGACGAAGATAC 1980
Db 1921 AACAGATACCTGCAACATCGCAATCAACGAGAGGCGCTTAAACATGACGAAGATAC 1980
Qy 1981 TATTAAGCTGTAAGAACTCAGAGACACAGACTCAGAAATACGTTCTGATTTAGAGCACA 2040
Db 1981 TATTAAGCTGTAAGAACTCAGAGACACAGACTCAGAAATACGTTCTGATTTAGAGCACA 2040
Qy 2041 AAGTGTATTAAGCGAGAAAGACGAGGTCCTTGTGCTTAAACCGGTGATCTGTAGATCA 2100
Db 2041 AAGTGTATTAAGCGAGAAAGACGAGGTCCTTGTGCTTAAACCGGTGATCTGTAGATCA 2100
Qy 2101 CCATTTACAGAGTTGGCTACAGAGTCTCAAGACACAGACCAAGAGACCTCACAAAGTC 2160
Db 2101 CCATTTACAGAGTTGGCTACAGAGTCTCAAGACACAGACCAAGAGACCTCACAAAGTC 2160
Qy 2161 CCAACCATCGGAGTCTATGAGTCCAGGTTCAAGTAACTGGAATCATCAAAAAGGCT 2220
Db 2161 CCAACCATCGGAGTCTATGAGTCCAGGTTCAAGTAACTGGAATCATCAAAAAGGCT 2220
Qy 2221 GTGACTAAGAAAGATCTGTTGTAGTGCAGAAAGGAAAACTGCGAGAAATCATCAGG 2280
Db 2221 GTGACTAAGAAAGATCTGTTGTAGTGCAGAAAGGAAAACTGCGAGAAATCATCAGG 2280

OY	2281	GATGTAAGAGAGATGAGACGTATGGAATGTTGCTGTAGACCTGTGATTCAAGTCTTCTTA	2340
Dp	2281	GATGTAAGAGAGATGAGACGTATGGAATGTTGCTGTAGACCTGTGATTCAAGTCTTCTTA	2340
OY	2341	AATGGGGTTAAGACACCCCGTTAAACCTCTGTATCTTGATGAGGACATTTGCGCCATGCA	2400
Dp	2341	AATGGGGTTAAGACACCCCGTTAAACCTCTGTATCTTGATGAGGACATTTGCGCCATGCA	2400
OY	2401	GGGACGCTGTGGACACTAATTCGATCGTCAAACTTAAGAAAGTGATATGTGCGGGAGC	2460
Dp	2401	GGGACGCTGTGGACACTAATTCGATCGTCAAACTTAAGAAAGTGATATGTGCGGGAGC	2460
OY	2461	CCAAAACATGCGGGCTTCTTTAAACATGATATGTCCTGAAAGTACATTTTAAACATATACATA	2520
Dp	2461	CCAAAACATGCGGGCTTCTTTAAACATGATATGTCCTGAAAGTACATTTTAAACATATACATA	2520
OY	2521	TGCACTGAAGTGAACCATTAATAAGACATCTCTAGAGAGTGCACACAGACTGTAAACCGCCATC	2580
Dp	2521	TGCACTGAAGTGAACCATTAATAAGACATCTCTAGAGAGTGCACACAGACTGTAAACCGCCATC	2580
OY	2581	GTTCTCCACGCTCTTTTACGACAAAGGAAATGAAGACGGTTAAACCATGTGCTGATTAATAATC	2640
Dp	2581	GTTCTCCACGCTCTTTTACGACAAAGGAAATGAAGACGGTTAAACCATGTGCTGATTAATAATC	2640
OY	2641	ATCATATAGATACCAACAGGGACCAAGAGCGGACAAAGATGATCTGATTTCTAACCTGTTTC	2700
Dp	2641	ATCATATAGATACCAACAGGGACCAAGAGCGGACAAAGATGATCTGATTTCTAACCTGTTTC	2700
OY	2701	AGAGATGGGTGAACACAGCTACAGATTGACTACAAATAATCAAGAAATCATGACTGCGGCT	2760
Dp	2701	AGAGATGGGTGAACACAGCTACAGATTGACTACAAATAATCAAGAAATCATGACTGCGGCT	2760
OY	2761	GCATTCGCAAGGACTTAACGCGGAAAGGCGTTTATGCTGTACAGTACAAAGTCAACAGAAAT	2820
Dp	2761	GCATTCGCAAGGACTTAACGCGGAAAGGCGTTTATGCTGTACAGTACAAAGTCAACAGAAAT	2820
OY	2821	CCACTCTACTCGACAGACTTTGAGACGCTGAACGCTTACTTACACGCAACGAAAGAAAGC	2880
Dp	2821	CCACTCTACTCGACAGACTTTGAGACGCTGAACGCTTACTTACACGCAACGAAAGAAAGC	2880
OY	2881	ATTGTCTGGAAGACGCTAGCTGTGTATCCCTGGATTAAGACACTTACAGCTAAATATCC	2940
Dp	2881	ATTGTCTGGAAGACGCTAGCTGTGTATCCCTGGATTAAGACACTTACAGCTAAATATCC	2940
OY	2941	GGGATTTTCAAGGCTTCAATTGAGACACTGCGACGCGGAAACAGACGCCATTTATGGCAGC	3000
Dp	2941	GGGATTTTCAAGGCTTCAATTGAGACACTGCGACGCGGAAACAGACGCCATTTATGGCAGC	3000
OY	3001	GTTCTGTGTAAGCCGACAGACTGATGTGTTCCAGAAATTAAGTGAACGTCGCTGGGCG	3060
Dp	3001	GTTCTGTGTAAGCCGACAGACTGATGTGTTCCAGAAATTAAGTGAACGTCGCTGGGCG	3060
OY	3061	AAGGCTTTAGAGCCAGTCTTGAGGACACGCGCAAACTTGTCTGACGACACAGAGTGGAG	3120
Dp	3061	AAGGCTTTAGAGCCAGTCTTGAGGACACGCGCAAACTTGTCTGACGACACAGAGTGGAG	3120
OY	3121	ACGTTGCAACCCATTCAGACATGACAGACGATCTACCTGAAATGAGCATTAACCTTCTTT	3180
Dp	3121	ACGTTGCAACCCATTCAGACATGACAGACGATCTACCTGAAATGAGCATTAACCTTCTTT	3180
OY	3181	TGCAACAGAGTCTTTGAGTAGACCTGACAGAGTGGGTAATTTTCCGCTCTTAACGCTGCAC	3240
Dp	3181	TGCAACAGAGTCTTTGAGTAGACCTGACAGAGTGGGTAATTTTCCGCTCTTAACGCTGCAC	3240
OY	3241	CTTACTTACAGAGGATCAGACCTGGGATATCTGCGCAAGGAAAGACATGTATGGGCTTAAT	3300
Dp	3241	CTTACTTACAGAGGATCAGACCTGGGATATCTGCGCAAGGAAAGACATGTATGGGCTTAAT	3300
OY	3301	AGAGAGGTAGCAAAAGAGTGTGTCACGGGAGATTCGCTGATCAACAAAGCGGTTGACACA	3360
Dp	3301	AGAGAGGTAGCAAAAGAGTGTGTCACGGGAGATTCGCTGATCAACAAAGCGGTTGACACA	3360

OY	3361	GGCAGGGTAGCGTATATTAAGAAATAATACATCAAGGACTACTCTCCAAACAATTAATGTC	3420
Db	3361	GGCAGGGTAGCTGATATTAAGAAATAATACATCAAGGACTACTCTCCAAACAATTAATGTC	3420
OY	3421	GTTCATTAATAATCCCGGTTGCCCACTCGTTGATCGTTGACCAACAAGACAGGGTACA	3480
Db	3421	GTTCATTAATAATCCCGGTTGCCCACTCGTTGATCGTTGACCAACAAGACAGGGTACA	3480
OY	3481	ACTGATCACAGCGGATTCCTATCTTAAGATGAAGGGCAAAATCTGTGGTATCGCGCAT	3540
Db	3481	ACTGATCACAGCGGATTCCTATCTTAAGATGAAGGGCAAAATCTGTGGTATCGCGCAT	3540
OY	3541	CCTATACGACATTCACGGGAAAGATAGATCCATGGGTCCATTGGGCCCACTAATACCATC	3600
Db	3541	CCTATACGACATTCACGGGAAAGATAGATCCATGGGTCCATTGGGCCCACTAATACCATC	3600
OY	3601	AGGTGTGATCTCGATTTTGGGAATACCTAGCCATGTCGGTAAATATATGACATTATCTTGTG	3660
Db	3601	AGGTGTGATCTCGATTTTGGGAATACCTAGCCATGTCGGTAAATATATGACATTATCTTGTG	3660
OY	3661	AATGTTAAGACCCCGTACAGGAACCATCACTACCAACGTCGAGGATCACGCTATCCAC	3720
Db	3661	AATGTTAAGACCCCGTACAGGAACCATCACTACCAACGTCGAGGATCACGCTATCCAC	3720
OY	3721	CACAGCAATGCTAATAGTGTAAAGGCTGTCCACCACTGGACGAACATGTGTGGCT	3780
Db	3721	CACAGCAATGCTAATAGTGTAAAGGCTGTCCACCACTGGACGAACATGTGTGGCT	3780
OY	3781	ATTAGGGTATGGGGCTTGGCTGATCGGCGCAACGAGAAATATCATCACTGCGGGTGGACGCTCA	3840
Db	3781	ATTAGGGTATGGGGCTTGGCTGATCGGCGCAACGAGAAATATCATCACTGCGGGTGGACGCTCA	3840
OY	3841	TTTAGGTTTACCCGTCGTCTGTCAAGCTTAAGAAACCTGCGGAAAAATCTGAAGTTCTCTTC	3900
Db	3841	TTTAGGTTTACCCGTCGTCTGTCAAGCTTAAGAAACCTGCGGAAAAATCTGAAGTTCTCTTC	3900
OY	3901	GTGTTCTTTCGGCAAGGACACGCAACCAACATGACCAAGACACTTGGGTAGTGTG	3960
Db	3901	GTGTTCTTTCGGCAAGGACACGCAACCAACATGACCAAGACACTTGGGTAGTGTG	3960
OY	3961	CTTACCAACATCTATCAAGGGTCAACACAGATAGAGGAGGGAGAGGCTCCAGCGTACAGA	4020
Db	3961	CTTACCAACATCTATCAAGGGTCAACACAGATAGAGGAGGGAGAGGCTCCAGCGTACAGA	4020
OY	4021	GTGATCAGAGGTGACATTAGCAAGAGGCGCTGACCAAGCTATCGTTAAATGTCTGTAATATGC	4080
Db	4021	GTGATCAGAGGTGACATTAGCAAGAGGCGCTGACCAAGCTATCGTTAAATGTCTGTAATATGC	4080
OY	4081	AAAGGTCAACCAAGTTCGAGGTGCGGTGCTCACTGTACCGAAAAATGGCGGCTGTTTT	4140
Db	4081	AAAGGTCAACCAAGTTCGAGGTGCGGTGCTCACTGTACCGAAAAATGGCGGCTGTTTT	4140
OY	4141	GATAGACAGCCAAATAGCTGTCTGGGACCGGCTTAGACTTGTGAAGACGAAACCGCTCATCTA	4200
Db	4141	GATAGACAGCCAAATAGCTGTCTGGGACCGGCTTAGACTTGTGAAGACGAAACCGCTCATCTA	4200
OY	4201	CATAGCTGAGGACCCAAATTTTCTTAAGATGTGCGGAAACCGAGGCGGACCTTAAGCTCGCA	4260
Db	4201	CATAGCTGAGGACCCAAATTTTCTTAAGATGTGCGGAAACCGAGGCGGACCTTAAGCTCGCA	4260
OY	4261	GCTGCTTACATAGACATAGCGTCCATCGTCAACCTGACGCGGATTAACAAAATATCAAGTA	4320
Db	4261	GCTGCTTACATAGACATAGCGTCCATCGTCAACCTGACGCGGATTAACAAAATATCAAGTA	4320
OY	4321	CCGCTACTGTCAACCGGATCTTATTTCTGGGTGGCAAAATCGAGTGAATCAATTCAT	4380
Db	4321	CCGCTACTGTCAACCGGATCTTATTTCTGGGTGGCAAAATCGAGTGAATCAATTCAT	4380
OY	4381	CACCTGTTCACTGTTTTCGACACTACGAGATGCGGATGTCAATATATTTGCTTGGATTA	4440
Db	4381	CACCTGTTCACTGTTTTCGACACTACGAGATGCGGATGTCAATATATTTGCTTGGATTA	4440
OY	4441	CAATGGAGACCAAGATTAATCGAGGCAATTCACCGCAAGAAAGACGTCGAATTTCTGGAT	4500

Db 4441 CAATGGAGACACGATATATGAGGCGATTCACCGCAAGAAAGCTCGAAATTCGTGAT 4500
QY 4501 GATGACAGGACGATGAGATTCATGATTCGAGGGTCCACCCAAACACTCTTTGGCAGGC 4560
Db 4501 GATGACAGGACGATGAGATTCATGATTCGAGGGTCCACCCAAACACTCTTTGGCAGGC 4560
QY 4561 AGACGAGGATTCCTGCTCAATGAGGAGGCAAGTTGATTCACTCTGAGAGTACAGATTC 4620
Db 4561 AGACGAGGATTCCTGCTCAATGAGGAGGCAAGTTGATTCACTCTGAGAGTACAGATTC 4620
QY 4621 CATGACAGCGGCAAGGACATTCGCGAAATCCATGCAATGTGGCCAACTGAGGCT 4680
Db 4621 CATGACAGCGGCAAGGACATTCGCGAAATCCATGCAATGTGGCCAACTGAGGCT 4680
QY 4681 AATGACGAGATTCCTGCTCAATGAGGAGGCAAGTTGATTCACTCTGAGAGTACAGATTC 4740
Db 4681 AATGACGAGATTCCTGCTCAATGAGGAGGCAAGTTGATTCACTCTGAGAGTACAGATTC 4740
QY 4741 CCAGTAGAGGATGAGAGGCGTCTGCTCACTCACTCACTCCATGCGCTGTGATTAAC 4800
Db 4741 CCAGTAGAGGATGAGAGGCGTCTGCTCACTCACTCACTCCATGCGCTGTGATTAAC 4800
QY 4801 GCTATGACGCGTGAAGCGGTATACAGGTTGCGCTCTGCGAAGAAAGACAGTTGCGGTA 4860
Db 4801 GCTATGACGCGTGAAGCGGTATACAGGTTGCGCTCTGCGAAGAAAGACAGTTGCGGTA 4860
QY 4861 TGCTCATCATCTCTGCTGCGCAAGTACAGATTCAGAGCGTGCAGAAAGTCACTGACG 4920
Db 4861 TGCTCATCATCTCTGCTGCGCAAGTACAGATTCAGAGCGTGCAGAAAGTCACTGACG 4920
QY 4921 AAACAGGCTCTGTTTTCAGGCGTGTACCAAGGCTGTGACACCCAGAGAAATGACGCGAA 4980
Db 4921 AAACAGGCTCTGTTTTCAGGCGTGTACCAAGGCTGTGACACCCAGAGAAATGACGCGAA 4980
QY 4981 ATAAATCTAGAAACGCCACC 5000
Db 4981 ATAAATCTAGAAACGCCACC 5000

RESULT 2
US-09-991-258-1
Sequence 1, Application US/09991258
Patent No. US20020141975A1
GENERAL INFORMATION:
APPLICANT: Olmsted, Robert
APPLICANT: Keith, Paula
APPLICANT: Dryga, Sergey
APPLICANT: Caley, Ian
APPLICANT: Maughan, Maureen
APPLICANT: Johnston, Robert
APPLICANT: Davis, Nancy
APPLICANT: Swanstrom, Ronald
TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOSES WITH MODIFIED HIV GENES FOR USE
FILE REFERENCE: 0113, 0001U3
CURRENT APPLICATION NUMBER: US/09/991,258
CURRENT FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: 09/902,537
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: 60/216,995
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 12523
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence; No. US20020141975A1 =
US-09-991-258-1

Query Match 42.3%; Score 2116.4; DB 10; Length 12523;
Best Local Similarity 64.5%; Pred. No. 0;
Matches 3218; Conservative 0; Mismatches 1746; Indels 24; Gaps 3;
QY 10 ACTAATGATTCATATATGAAAGAAATTCACGTTGACTTAATGCTGACAGCCCGTATGTC 69
Db 30 AATTAATCTACCCAAATGGAAGAAAGTTCAAGTTGACATCGAGAAAGACACCAATTCCTC 89
QY 70 AAGTCGTTACAGGAGGAGCTTCCACAATTTAGATGGAAGCAAGGAGGCTGACGCAAT 129
Db 90 AGAGCTTTCAGAGGAGAGCTTCCGCACTTTGAGTAGAAGCCAAAGAGTCACTGATTAAT 149
QY 130 GACCATGCAATGCAAGAGGCTTTGCGATGTCGCAACAAAGCTCAATTGAGAGGAGTGC 189
Db 150 GACCATGCTAATGCGAGAGGCTTTGCGATGTCGCAACAAAGCTCAATTGAGAGGAGTGC 209
QY 190 GACCGGAGCAAGTATCTTGGACATTTGGAAGTGGCCGCTGAGACATGCAATTCAT 249
Db 210 GACCATGCGACAGATCTTGGACATTTGGAAGTGGCCGCTGAGACATGCAATTCAT 269
QY 250 CACCGTATGATTTGATCTGCGCTATGATGATGAGGCTGGAAGCCGAGACATACAAAGG 309
Db 270 CACAGTATCATTTGATCTGCGCTATGATGATGAGGCTGGAAGCCGAGACATACAAAGG 329
QY 310 TATGCAAGAAAGCTTAAGAAAGT-----GACATTAACGACAAAGACATAGCTCTTAAG 363
Db 330 TATGCAAGTAAAGCTTAAGAAAGTAACTGTAAGAAATTAATGTAAGAAATTTGCAAGAAA 389
QY 364 GCGGAGAGCTGCTGGAAGTATGTCACACGAGAGGAGAGTCCATCTCTGTATG 423
Db 390 ATGAGAGAGCTGCGCGCTGATGAGGAGCCTGAGACCTGGAACCTGAGATATGTGCTC 449
QY 424 CACACAGGCGCAGTGTAGTATCTTGGAAAGTGTAGCAATATACCAAGATGTGTAGCA 483
Db 450 CACAGCAGAGTGTGTGCTGCTAGAGAGGAGGAGTGTGTATACAGATGTGTAGCG 509
QY 484 GTCCATGCAACGACATCAATCTACCAAGGCGCTTAAGAGTGTAGCAATTTACTG 543
Db 510 GTTACGAGAGCGAAGTGTCTATACCAAGGCGCTTAAGAGTGTAGCAATTTACTG 569
QY 544 ATAGGCTTTGACACGAGCCCTTTTATGTACAAAACATGCGAGTTCTACCTACTTAC 603
Db 570 ATAGGCTTTGACACGAGCCCTTTTATGTGTAGAAAGTGTGAGCATTCATCTATAC 629
QY 604 AACAGAACTGGGCTGAGAGAGATTTGGAAGACATGTAACATTTGCGTCTGTAATCA 663
Db 630 TCTACACAGTGGGCGAGAGAACCGTGTAAAGGCTGTGTAACATGAGCTATGAGCTCT 689
QY 664 GATCTTCAGAGAGAGGCTTTGGAAGAACTCTCAATCTTGAAGAGAGGCTTCAACT 723
Db 690 GAGCTTATGAGAGGCTGAGGTAGAGGATGCTCAATCTTGAAGAGAGTGTGAAACCA 749
QY 724 ACTAATGATCATATTTCTGCTGTTGTTCAACATCTACACAGAGATGATCACTGTTA 783
Db 750 TCCAAATATTTCTATCTGTTGCTGACCAATCTACACAGAGAGAGGAGGACTTACTG 809
QY 784 CTAAGCTGAGATCTTCCAAAGCTGTTCCACTGTAAGAGAGAGTCACTTCAAGGTAGA 843
Db 810 AGAGCTGAGACCTGCGCTGTATTTCACTTACGTGCAAGCAAGAAATTAACACATGTCG 869
QY 844 TGTGAGACATTTGTCAGCTGTAAGGAGTACGTATCAAAAGATTAACATGAGCCGAGCA 903
Db 870 TGTGAGACATTTGTCAGCTGTAAGGAGTACGTATCAAAAGATTAACATGAGCCGAG 929
QY 904 CTATACGTTAAAGTTGAACTTGGCGTCAACATGATGCGAGGCTTTCTGAGTTGC 963
Db 930 CTGATGAGAAAGCTTCAAGCTATGCTGATACATGACGCGAGGAGTCTTGTGCTGC 989
QY 964 AAAGTCAAGATACGCTGCGCGAGAGGTTCTTCTGTGTAGTATGATGCA 1023
Db 990 AAAGTCAAGATACGCTGCGCGAGAGGTTCTTCTGTGTAGTATGATGCA 1049
QY 1024 GCCACATTTGCGATGATGATGACAGGATTCGGAACATGACGTTAGTGTAGTACGCA 1083

D	b	1050	GCTACATTGTGTGACCAATATGACTGCGTACTCGGACACAGATGTGACGTGCGGACGACCGG	1109
Q	y	1084	CAAAAACCTATTGGTTGGGCTCAACCCMAAGATTGTGCTCAATGGTAGGACGCCAAGAAAT	1143
D	b	1110	CAAAAACCTGCTGGTTGGGCTCAACGAGGTATATGTCGTCAACGGTCGACCCAGAAAC	1165
Q	y	1144	ACTAACACAAATGCAACAATCTATCTATTTACAGAGTGTGTCGCCAGGCGCTTTTCCAGGTGGCG	1203
D	b	1170	ACCAATACCAATGAAAATTTACCTTTTGGCCGTATGGGCCACAGGCACTTTGTCTAGTGGGCA	1229
Q	y	1204	CGTGAACATCGTGCCGCACTTGGACGAGAGAAAGAACTATGGGGTGGGGAACGCACTTT	1263
D	b	1230	AAGGATATTGAAGAAATCAAGAAATGAAAGCCACTAGACCTACGAGATGACAGTTA	1289
Q	y	1264	ACTATGGGCTGCTGCTGGCTTTCCAGAACCCAGAAATCACATCATCTTACAGAAGCCT	1323
D	b	1290	GTCAATGGGGTGTGTGGGCTTTTAAAGGACAAAGATATCTATTTATTAAGGCCCGC	1349
Q	y	1324	GGTACGCAACAAATTAGAAGATACCTGCCCTCTTTGACTCATTTGTGATTTCCAGCCCTT	1383
D	b	1350	GATACCCCAACCATCATCATCAAAAGTGAACGCAATTTCCATCTATCTGTCCTCCAGATTA	1409
Q	y	1384	ACCAAGCCACGGGCTCGATATGGGCTTCCGCGTAGCTCAAGCTGCTCTGAACCAACT	1443
D	b	1410	GGCAGTAAACATTTGAGATGCGGCTGAGAACAAAGATCAGAAATTTTAGAGAGCAC	1469
Q	y	1444	GTCAAAACCCGACCGGCTATTATCAATAGCCGATGTGAGCAATCTGCGGTACTACAGCA	1503
D	b	1470	AAGGAGCCGTACACTCTCTCATTTACCCGCGAGAGACGTACAAAGAACTATAGTGGCACCGAT	1529
Q	y	1504	GAAGCTGAAGAAATGTCGCTGACAGCCGAGAGATCAGAGAAAGCCCTGCCACCTTGCCT	1563
D	b	1530	GAGCTTAAGGAGGTGCGTGAAGACGAGAGATTGGCGCGCACTTACCACTTTGGACAGCT	1589
Q	y	1564	GAATTTAGAAAAGAGACCGGTAGAGGCAAGATGAACTCTATTATGCAAGAGGACGAGACA	1623
D	b	1590	GATGTGAGAGGCCCACTCTGAAAGCCGATGTCACTTGAATGTTACAAAGAGCTGGGGCC	1649
Q	y	1624	GGTAGCGTGAGACACCAAGAGACCACTACAGGGTGAACAAGTTTACCAGGCGGAAGAGAG	1683
D	b	1650	GGCTCAGTGGAGACACCTCGTGGCTGTGATTAAGGTATCCAGCTACGCTGGCGAAGACAG	1709
Q	y	1684	ATTGGGCTTACGCTATATCTTTCAACCCAGGCGGTTGAATAGTGAATAAATCGGCGTGT	1743
D	b	1710	ATCGGCTCTTACGCTGTCTTTCTCCGAGGCTGTACTCAAGAGTGAAATAATTTACTTGC	1769
Q	y	1744	ATCCACCCATTGGGGGAAACAGTACTGTGTAATGACTCAAAAGTATGGGCAAGGAGATAC	1803
D	b	1770	ATCAACCCCTCTCGCTGAACAAGTCAATATGATTAACAACCTGCGCCGAAAAGGCGCTTAT	1829
Q	y	1804	AAAGTCAGACATTAACCAAGTAAAGGTCACTTGAACAGAAAGGACGCGCGTCCCTGTTCAA	1863
D	b	1830	GCGGTGAACCATTAACATGATTAAGTATGTGTGTCGAGGGAATGCAATACCCGTCGAG	1889
Q	y	1864	GACTTCCAGGCAATTGATGAGAGGCGTACGATCGTTTTCACAGAGGGAATTTGCTAAAC	1923
D	b	1890	GACTTTCAAGCTCTGAAGTGAAGGCCACCAATTGTGTACAACGAACGTGAATTTGCTAAAC	1949
Q	y	1924	AGATTAAGTCAACCAATGCAATCAACGAGAGAGCGCTTAAACCTGACGAAGAGTACTAT	1983
D	b	1950	AGGTACCTGCAACATATTGCCACACATGAGAGAGCGCTGAACACTGTATGAAGAAATATTAC	2009
Q	y	1984	AAGACTGTAAGACATCAAGGACACAGACCTCAAAATACGCTTCGATTTGAAGCAGCAAGAG	2043
D	b	2010	AAAACTGTCAAGCCGACGCAACGACGCGGAATATCTGTACGACATTCAGACGAGAAACAG	2069
Q	y	2044	TGTGTTAAGCAGAGAAAGCAGAGTCCCTTGTGCTTAACCGGTGATCTGTGATCTCAACA	2103
D	b	2070	TGCGTCAAGAAAGAACTATGTCACCTGGGCTTAAGGGCTCAACGCGAGAGTGTGTGATCTCC	2129
Q	y	2104	TTTCAACGAGTTTGGGTATGAGAGTCTCAAGACAGACGACGACCTCACAAAAGTCCCA	2163

Db	2130	TTCCATGAAATTCCGCTTACGAGAGTCTGAGAACACGACACGACCGCTCCTTACCAAGTACCA	2189
Qy	2164	ACCATCGGAGTCTATGAGAGTCCAGGTTCAAGTAAATCTGGAAATCATCAAAAGCGCTGTG	2223
Db	2190	ACCATAGGGGGGTATGGGGGTGCGAGGATCAGGGCAAGCTGGCATCTTAAAGGCGAGTC	2249
Qy	2224	ACTAAGAAAGATCTGGTTGTGAGTGGCGAAGAGAAAACCTGCCGAAATCATAGGGAT	2283
Db	2250	ACCAAAAAGATCTAGTGGTGAAGGCCAAGAAAACCTGTGCAGAAATTTATMAAGGAC	2309
Qy	2284	GTAAGGAGGATGAGACGTAATGGATGTTGCTGTGAGCTGTGCATTCAGTGCCTCTAAAT	2343
Db	2310	GTCAAGAAATGAAAGGCGTGGAGCTCAATGCCAAACTGTGGCTAGCTCAGTGCCTTAAAT	2369
Qy	2344	GGGGTTAAGACACCCCGTTAACCTCTGTACATTGATGAGCATTTGCCCTGCATGACAGG	2403
Db	2370	GGATGCAAAACACCCCGTAGAGACCCTGTATATTTAGCAAGCTTTTGGTTGTCAATGACAGT	2429
Qy	2404	ACGCTGCTGGCATGTGATTTGCCATTCGTCAAACTTAAGAAAGTGGTATTTGGGGGACCA	2463
Db	2430	ACTCTCAGAGGCGCTCATAGCCATTATTAAGACTTAAAAAGCAGTGCCTCGGGGGATCCC	2489
Qy	2464	AAACAAATGGCGCTCTTTAACATGATGTCGCTGAABATACATTTTAAACATGACATATGC	2523
Db	2490	AAACAGTGGGTTTTTTTAAACATGATGTGCTTGAAATGCAATTTTAAACACAGATTTGC	2549
Qy	2524	ACTGAGGTATCACTAAAGACATCTCTAGAGAGTGCACACAGACTGTAAACCGCATGCTC	2583
Db	2550	ACACAGTCTTCCACAAAGATCTCTCGCGTTGCACCTAAATCTGTGACTTGGTGTGTC	2609
Qy	2584	TCCACGCTCTTTCATGCAACAGCGAATGAAGCGGTTAACCCATGTGCTGATTAATCATC	2643
Db	2610	TCAACCTGTGTTTTCGACAAAAAAATATAGAACGACGAATCCGAAAGAGACTAAGATTGTG	2669
Qy	2644	ATAATATCCACAGGAGCACAAAGCCGACAAAGATATCTGATTTCTAACTGTGTTTACA	2703
Db	2670	ATTACACTACCGGACATGACCAACCTTAGAGAGCAGATCTCATTTCTACCTTGTGTTACA	2729
Qy	2704	GGATGGGTGAACAGCTACAGATTGATTAACAAAAATCAGCAAAATCATGACCTGCGGCTGCA	2763
Db	2730	GGGTGGGTGAAGCACTTGGCAATATGATTAACAAAGGACGAAATATATGACGGCAGCTGCC	2789
Qy	2764	TGCGAAGGACTTACGCGGAAAAGCGGTTATGTCTGTCAAGTACCAAGTCAACGAAATCCA	2823
Db	2790	TCTCAAGGGCTGACCCCGTAAAGGGTGTATGCGGTCGGTACAAAGGTAATGAAAAATCCT	2849
Qy	2824	CTCTAATCTCGCAGATTTCTGAGCAGCTGAACGTGTACTTACAGCCACAGAAAAACGATT	2883
Db	2850	CTGTAGCAGCCACCTCAGAAACATGTAAAGTCTCTACTGACCCGACAGAGAGACCCGATC	2909
Qy	2884	GTCGTGAAGAGACGTAGCTGTGTGATCCCTGTGATTAAGACCTTACAGCTAAATATCCGGG	2943
Db	2910	GTGTGAAAAACATAGCCGGCGACCCATGTATTAACACTGACTGCGCAAGTACCTCTGGG	2969
Qy	2944	GATTTCACGGCTTCATTGAGCAGACTGTGCAGCGGACGCAACGACGCGCATTTATGACGCGTT	3003
Db	2970	AATTTCACTGCAAGATAGAGAGGTGCAAGCAGAGCATATGTGCATCATGAGGCAATC	3029
Qy	3004	CTTGTATTAACCCGACAGACGCTGATGTGTTCCAGATTAAGTGAACTGTCTGCTGGCGAAG	3063
Db	3030	TTGAGAGAGACCGGACCTTACCGAGTCTTCAGATTAAGGCAAAACGTGTCTGGGCGCAAG	3089
Qy	3064	GCTTTAGGCGAGCTTGGGCAAGGGCCAAACATTTGTGTCGAGACAGCAGCTGGGAGACG	3123
Db	3090	GCTTTAGTCCGGGTCTGAAGAACCGCTGGCATACATGACCACTGAACATGGAACACT	3149
Qy	3124	TTGCACCCATTCACAGCATGACAGACGCTACTCACTTGAAATGGCACTGAATCTTTTTCG	3183
Db	3150	GTGATATTTATTTGAAACGGAACAAAGCTCACTCAGCAGAGATAGTATTTGAACCAACTATGC	3209
Qy	3184	ACACAGGTTCTTTGGAGTAGACCTGTGACAGATGGGGTATTTTTCGCTCTTACCTTCGCACCT	3243
Db	3210	GTGAGGTTCTTTGGACATCGATCTGGACATCGGGCTATTTTCTGACACCCATGTTTCGGTTA	3269

QY 3244 ACTTACAGGATCAGACATGCGGATTAATCGCCAGGAAAGATGATGAGCTTAATAGA 3303
 Db 3270 TCCATTAGGAATATCACTGGGATTAATCCCGCTCGCTTAATGATGAGGCTTAATAGA 3329
 QY 3304 GAGGTACCAAGAGGATGTCAGGCGATATCGTGCATCAACAAAGCGGTTGACACAGGC 3363
 Db 3330 GAATGTGTCGCTGACTCTCGCAGGTAACCAACATGCTCGGGCAGTTGGCCATGGA 3389
 QY 3364 AGGTACTGATTAATAGGAATTAATCAATCAAGATTAATCTCTCAACATTAATGTGTT 3423
 Db 3390 AGAATCTATGACATGAACACTGTAACACTGCGCAATTAATGATCCGCAATTAACCTGTA 3449
 QY 3424 CCATTAATCGCGGCTGGCCCACTGCTGATCGTGAACCAAGGACAGGATACAAT 3483
 Db 3450 CTTGTAAGCAAGAACTGCTCAATGCTTAATCTTCAACCAATTAATGAACCCACAGAGT 3509
 QY 3484 GATCAGACGCGATTCCTATCTAAGATGAAGGCAAAATCTGTGTGTGATCGGCGATCCT 3543
 Db 3510 GACTTTCTTCTATTCGTACGCAAAATGGAAGGGCAAGAACTGCTGTGTGTGCGGGAAAAAG 3569
 QY 3544 ATCAGCATTCAGGAGGAAGATGAGTCATGAGTCCATTTGCCCATTAATCCATCAGG 3603
 Db 3570 TTGTCCGTCAGGCAAAATGTTGACTGTGTGACAGCCGCTGAGGCTTAACCTTGA 3629
 QY 3604 TGTGATCTCGATTTGGGAAATCACTAGCCATGTGCTGTAATTAATGACATTAATTTGTCAAT 3663
 Db 3630 GCTGGCTGATTTAGGATTCAGGATCCAGGTGATGTGCCCAATTAATGACATTAATTTGTTAAT 3689
 QY 3664 GTTAGAACCCCGTACAGGAACCATCACTACCAACAGTGCAGAGATCAAGCTATCCACAC 3723
 Db 3690 GTGAGGACCCCATTAATTAATCACTACATCAGAGTGTGAAGACCAATGACATTAATGCTT 3749
 QY 3724 AGCATGCTAAGTGAAGCTGTGTCACCACTGAACCTGCGGAAACATGTGTGCTATA 3783
 Db 3750 AGCATGTTGACAAAGAACTGTGTGATCTGAATCCGCGGAACTGTGTGACGATA 3809
 QY 3784 GGGATGGGCTTGTGATCGCGAACCGAATATCACTGCGGTGACAGCTCAATTT 3843
 Db 3810 GGTATGTTAGCTGACAGGCGCAAGCAATCACTGTGTCTATTAAGCGGCGAGTTC 3869
 QY 3844 AGGTTAACCCGCTGTCTGACCTTAAGAACATGCGGAAATTAAGTGAAGTTCTTCTGTG 3903
 Db 3870 AAGTTTCCGCGGATGCAACCGAAATCTCACTGAAGAGACGAAATTTCTGTTGTA 3929
 QY 3904 TTCTTCGCAAGGACCAACGCAACCAATATCAACAGGACAGTCTGCTGTGTGCTT 3963
 Db 3930 TTCAATGGGATCGATCGAAGGCGCTGTAACCAATCTTCAAGCTTTCATCAACCTTG 3989
 QY 3964 GACAACTCTATCAAGGCTCAACAGGTAAGGACAGGAGCTCCAGCTGATCAGAGTG 4023
 Db 3990 ACCAACTTTATTAAGTTTCCAGACTCCAGCAAGCCGATGTGACCTCATATCATGTG 4049
 QY 4024 ATCAGAGGTGACATTAAGCAAGCGCTGACCAAGCTATGTTAATGCTGTAATAGCAA 4083
 Db 4050 GTGCGAAGGGGATTTGCGACGCGACCGAAGAGTGAATTAATTAATGCTGTAACAGCAA 4109
 QY 4084 GGTAAACAGGTTCCGAGTGTGGGTGCACTGTACGAAATATGGCGGCTGCTTTGAT 4143
 Db 4110 GGAACACCTGCGGAGGGGTGTGCGAGCGCTGTAATGAAGTTTCCGGAAGGCTTGAT 4169
 QY 4144 AGAAGCAATAGTGTGCGAGCGCTGACTTGTGAAGCAGCAAC---GCTCATATA 4200
 Db 4170 TTAAAGCCGATCGAAGTGAAGAAAGCGGACTGTGTAAGGTGACGTTAAACATATCAT 4229
 QY 4201 CATGCTATAGACCAATTTTCTAAGATGCGGAACCGAGGCGACCTTAAGCTGCA 4260
 Db 4230 CATGCTATAGACCAATTTTCTAAGATGCGGAACCGAGGCGACCTTAAGCTGCA 4289
 QY 4261 GCTGCTATAGACATAGCTCATGCTCAACGCTGAGCGGATTAACAAATATCACTA 4320
 Db 4290 GAGGCTATAGATCATGCTCAAGATGCTCAACATTAATCAAGTGTAGTCAAT 4349

QY 4321 CCGCTACTGTCAACCGGACATCTATTCTGTGTGGCAAGATGATGATCAATTCAT 4380
 Db 4350 CCACTGTGTCAACCGGACATCTTTCCGGAAACAAAGATGATGATCAATTCAT 4409
 QY 4381 CACTGTTCACTGTTTCCGACTACGATGCGGATGCGGATGCTATATTTGTTGATTA 4440
 Db 4410 CATTGCTGACAGTTTGAACCACTGATGATGATGATGATGATGATGATGATGATGAT 4469
 QY 4441 CAATGGGAGACGAGATTAATGAGGATTCACCGCAAGAAAGCGGCAATTTCTGA - 4499
 Db 4470 AATGGGAAATGACTCTCAAGAAAGAGCTGCTAAGAGAGAGAGAGAGAGATATGC 4529
 QY 4500 -----TGATGAACAAGCCAGTGAACATTCATGCTGTGAGGTTCAACCAAC 4545
 Db 4530 ATATCCGAGACTCTTCAAGTGAAGAGACCTGATGAGAGAGAGAGAGATATGC 4589
 QY 4546 AGCTTTTGGCAGGACGACAGGTTACTCCGTCAATGAGGCAAGTTGATTTACTG 4605
 Db 4590 AGTTCTTGGCTGGAAGAGGCTACAGCACAAGCGATGCAAACTTCTCATATTTG 4649
 QY 4606 GAAGGTACAGATTCATCAGACCGGCAAGACATTCGCAAAATCCATGCAATGTGCCC 4665
 Db 4650 GAGGAGCAAGTTTCAACGAGGCGCAAGAGATTAAGCAAAATTAATGCTCATGTGCC 4709
 QY 4666 AACAAATCTGAGCTAATGAGAGATTTGCTGTACATCTGCGGAGAGATGTCCAGC 4725
 Db 4710 GTTGCAACGAGGCAATGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 4769
 QY 4726 ATCCGCTCCAAATGCGGATGAGAGATGAGAGGCTGTCTCTCACTCAACACTTCCA 4785
 Db 4770 ATTAGGTGCAAAATGCGGCTGCAAGAGTGTGAAGCTGTCCACACCACTTACAGCTGCT 4829
 QY 4786 TGCTGTGTAATTAAGCTATGAGCGGCTGAGCGGCTATACAGGTTGCGCTCTGCAAGAA 4845
 Db 4830 TGCTGTGTAATTAAGCTATGAGCGGCTGAGCGGCTATACAGGTTGCGCTCTGCAAGAA 4889
 QY 4846 GAACAGTTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4905
 Db 4890 GAACAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4949
 QY 4906 AAGCTACAGTGCAGCAACAGATCTGTTTTCAGGCGCTGTACACGCGTGTACACCCC 4965
 Db 4950 AAGTCCAAATGCTCCGAGCTATATGTTCTCAACCGAAAGTGTGCTGTATATTCATCA 5009
 QY 4966 AGGAAGTA 4973
 Db 5010 AGGAAGTA 5017

RESULT 3
 US-09-991-258-14
 : Sequence 14, Application US/09991258
 : Patent No. US20020141975A1
 : GENERAL INFORMATION:
 : APPLICANT: Olmsted, Robert
 : APPLICANT: Keith, Paula
 : APPLICANT: Dryga, Sergey
 : APPLICANT: Caley, Ian
 : APPLICANT: Maughan, Maureen
 : APPLICANT: Johnston, Robert
 : APPLICANT: Davis, Nancy
 : APPLICANT: Swanson, Ronald
 : TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOSES WITH MODIFIED HIV GENES FOR USE
 : FILE REFERENCE: 01113.000103
 : CURRENT APPLICATION NUMBER: US/09/991.258
 : CURRENT FILING DATE: 2001-11-16
 : PRIOR APPLICATION NUMBER: 09/902,537
 : PRIOR FILING DATE: 2001-07-09
 : PRIOR APPLICATION NUMBER: 60/216,995
 : PRIOR FILING DATE: 2000-07-07
 : NUMBER OF SEQ ID NOS: 19
 : SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 14
; LENGTH: 12379
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. US20020141975A1e =
US-09-991-258-14

Query Match 42.3%; Score 2113.2; DB 10; Length 12379;
Best Local Similarity 64.5%; Pred. No. 0;
Matches 3316; Conservative 0; Mismatches 1748; Indels 24; Gaps 3;

QY 10 ACTAATCCATCCATATGGAAGAAATTCACGTTGACTTAGATGTCAGCCCGTATGTC 69
DB 30 AATTACCTACCAAAATGGAAGAAATTCACGTTGACTTAGATGTCAGCCCGTATGTC 89
QY 70 AAGTCGTTACAGCGGACGTTTCCACAAATTGAGATCGAAGCAAGGCAAGTCATGACAA 129
DB 90 AAGCTTTGACAGCGGACGTTTCCGCAAGTTGAGTAGAAGCCAGCAAGTCATGATAT 149
QY 130 GACCATGCCAATGCGCAGCGCTTTTCGATGTCAGCAACAAAGCTCATTGAGCGAAGTC 189
DB 150 GACCATGCTAATGCGCAGCGCTTTTCGATGTCAGCAACAAAGCTCATTGAGCGAAGTC 209
QY 190 GACCGGACCAAGTTATCTTGACATTTGGAAGTGGCCCGCTCAGACATGTCATTCAT 249
DB 210 GACCATCCGACAGCATCTTTCGATGTCAGCAACAAAGCTCATTGAGCGAAGTC 269
QY 250 CACCGCTATCATTTGATCTGCTTATGATTAAGCGCTGAAGACCCGGAACAGTACAA 309
DB 270 CACAAGTATCATTTGATCTGCTTATGATTAAGCGCTGAAGACCCGGAACAGTACAA 329
QY 310 TATGAGAAAGACTTAAGAAAGT-----GACATTAACGACAAAGCAATGACTCTTA 363
DB 330 TATGAGAAAGACTTAAGAAAGT-----GACATTAACGACAAAGCAATGACTCTTA 389
QY 364 GCGGAGACCTGCTGGAAGTCAATGTCACACAGGCGAGAGACTCCATCTGCTGATG 423
DB 390 ATGAGAGAGCTGCGCCGCTCATGAGCGACCTGACCTGGAACCTGAGACTATGCTG 449
QY 424 CACACAGACGCGACGTTGAGTACTTTGGAAGTGAAGCAATGACCAAGTGTACGA 483
DB 450 CACGACGAGAGCTGCTGCTGCTACGAGGCGCAAGTGGCTTTACAGAGTGTATAC 509
QY 484 GTCCATGACAGCAATCATCTACCAAGGCGCTTAAAGAGTGAAGCAATTTACTG 543
DB 510 GTTACGAGACGCAAGTCTTATCAACCAAGCAATGAAGGAGTTAGAGTCCCTACTG 569
QY 544 ATAGGCTTTGACAGACCGCTTTTATGTAACAAACCATGGAGGTTCCCTACCTTAC 603
DB 570 ATAGGCTTTGACAGACCGCTTTTATGTAACAAACCATGGAGGTTCCCTACCTTAC 629
QY 604 AACAGCACTGGGCTGACGAGAGATTTGGAAGCAGTAACTATGAGCTCGCTACTCA 663
DB 630 TCTACCACTGGGCGGAGGAAACGTTGTAAGCGCTCGTAACTATGAGCTCGCT 689
QY 664 GATCTTCAGAGAGAGGCTTGAAGAACTCTCAATCTTTAGAGAGAGAGGCTCCAACT 723
DB 690 GACGTTTGAAGAGGCTGACGTAAGAGGATGTCATTTGAAAGAAAGATTTGAAACA 749
QY 724 ACTAATAGATCATATTTCTGGTTGGTTCAACATCTACAGAGAGATGATCACTGTTA 783
DB 750 TCCAACTATGTTCTATTTCTGTTGCTGCTGACCACTACACAGAGAGAGGAGCTTACTG 809
QY 784 CGTAGCTGACATTTTCCAAACGTTGTTCCACTTGAAGGAAAGTCTTACTTACAGAGT 843
DB 810 AGGAGCTGACACTGCTGCTGTTATTTCACTTACGTTGAGCAAAATTTACAGTGG 869
QY 844 TGTGGAGCACTTGTGAGCTGTAAGGTAAGCTATCAAAAGATTAAGTACAGCCAGGA 903
DB 870 TGTGAGACTATATGTTGAGCTGTAAGGTAAGCTATCAAAAGATTTAGCTACGTCAG 929

QY 904 CTATACGTTAAAGTTGAGAACTTGGCTCCACAAATGATCCGAGGCTTTCTTGAAGTTC 963
DB 930 CTGTATGGGAAGCCCTTACGAGCTATGCTCTGATGATGACCCGGAAGGATTTCTGTGCTGC 989
QY 964 AAAGTCACAGATACCGCTCGCGGCGAGAGGCTTTTCTTGTGTGTATGATGATGACCA 1023
DB 990 AAAGTCACAGATACCGCTCGCGGCGAGAGGCTTTTCTTGTGTGTATGATGATGACCA 1049
QY 1024 GCCACACTTGGGATCAGATGACAGGAGTTCGCGCACTGACGTTAGTGTGATGACGCA 1083
DB 1050 GCTTACATTTGTTGACCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1109
QY 1084 CAAAAAATTTGTTGGCTCAACCAAGAGATTTGCTCAATGTTGAGACGCAAGAAAT 1143
DB 1110 CAAAAAATTTGTTGGCTCAACCAAGAGATTTGCTCAATGTTGAGACGCAAGAAAT 1169
QY 1144 ACTAATGATGACAGAACTATCTATTTACCAAGTGTGCTGCTGCTGCTGCTGCTGCTG 1203
DB 1170 ACAAATACCAATGAAAAATTTACCTTTTGCCTGATGCTGCTGCTGCTGCTGCTGCTG 1229
QY 1204 CGTGAACATCGTCCGACTTGGACGACGAAAGAACTTAAAGGCTGCGGAGCGCACTCT 1263
DB 1230 AAGGAATTTAAAGAAATCAAGAAATGAAAGCCACTAGACCTACGAGATGACAGTTA 1289
QY 1264 ACTATGCGCTGCTGCTGCTTTCAGAGCCGAAATACATCCATCTACAAAGGCT 1323
DB 1290 GTCATGCGCTGCTGCTGCTTTCAGAGCCGAAATACATCCATCTACAAAGGCT 1349
QY 1324 GGTACGCAACAAATTAAGAAAGTACCTGCTTTTGAATGATTTGATTCAGGCTT 1383
DB 1350 GATACCCCAACATCATCAAGTGAACAGCATTTCCACTATGCTGCTGCTGCTGCTGCTG 1409
QY 1384 ACCGACCAAGGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1443
DB 1410 GGCAGTAAACATTTGAGATGAGTGGCTGAGAACAAATGAGAAATTTTGAAGAGAC 1469
QY 1444 GTCAAAACCGGACCGGCTATTTACAAATGCGATGAGAGCACTGCTGCTGCTGCTGCT 1503
DB 1470 AAGGAGCGCTACCTCTCTATTTACCGCGAGAGAGCTACAAAGCTTAAGTGGCAG 1529
QY 1504 GAAGCTGAAGAGTGGCTGACGCGGAGAGATCAGAGAAAGCTTGCACCTTCTGCTCT 1563
DB 1530 GAGCGTGAAGAGTGGCTGACGCGGAGAGATCAGAGAAAGCTTGCACCTTCTGCTCT 1589
QY 1564 GAATTAAGAAAGAGACCGTGAAGGCAAGTGAACCTCATTTATGACAGAGGAGAGCA 1623
DB 1590 GATGTTGAGAGCCCACTCTGGAAGCGATGCTGATGATTTAAGAGGCTGGGCTC 1649
QY 1624 GGTAGCGTGAAGACCAACAGAGACATTCAGGAGTGAACATTTACCCGAGCAAGAGAG 1683
DB 1650 GGTAGCGTGAAGACCAACAGAGACATTCAGGAGTGAACATTTACCCGAGCAAGAGAG 1709
QY 1684 ATTGGCTTTACGCTATATCTTTCAACCCGAGCGGATTTGAATAGTGAAGAAATGCG 1743
DB 1710 ATTGGCTTTACGCTATATCTTTCAACCCGAGCGGATTTGAATAGTGAAGAAATGCG 1769
QY 1744 ATTCACCCATTTGGGAGAAAGTACTGATTAATGATCAACAAAGTGAAGGAGAGATAC 1803
DB 1770 ATTCACCCATTTGGGAGAAAGTACTGATTAATGATCAACAAAGTGAAGGAGAGATAC 1829
QY 1804 AAAGTCAGGACATCAAGGATGATTTGATGACAGAGAGGAGCGGCTGCTTCA 1863
DB 1830 GCGTGAACATATCAATGATGATTTGATGACAGAGAGGAGCGGCTGCTTCA 1889
QY 1864 GACTTCAGGATTTGATGATGAGAGCGCTTACGATGCTTTTCAACGAGAGGAGTCTTAAAC 1923
DB 1890 GACTTCAGGATTTGATGATGAGAGCGCTTACGATGCTTTTCAACGAGAGGAGTCTTAAAC 1949
QY 1924 AGATTAACCTGACCAATGATGACATCAAGGAGAGCGCTTAAACATGACAGAGAGATAT 1983
DB 1950 AGGTAACCTGACCAATGATGACATCAAGGAGAGCGCTTAAACATGACAGAGAGATAT 2009
QY 1984 AAGACTGTAAGAGCTCAAGGACAGAGACTCAAGATACGTTCTGATATTTGACGCAAGAG 2043

Db 2010 AAAAGTGTCAAGCCAGGAGACAGAGCGGAAATACCTGTACGATGACAGGAAACAG 2069
Qy 2044 TGTGTTAAGCGAAGAAAGACGAGGTCCCTTGTGCTTAACCGGTGATCTGTGTGATTCACCA 2103
Db 2070 TGCCCTCAAGAAAGAACTAAGTACTGGGCTAGGGCTCAAGGCGAGCTGTGTGATTCCTCC 2129
Qy 2104 TTTACGAGTTTGGGTACGAGAGTCTCAAGACGAGCAGCAGACCTCAAAAGTCCCA 2163
Db 2130 TTCCATGATTTCCCTTACGAGAGTCTGAAACAGACAGCCGCTCTTACCAAGTACCA 2189
Qy 2164 ACCATCGAGTCTATGAGTGCAGGTTCAAGTAATCTGAAATCATCAAAAGCGCTGTG 2223
Db 2190 ACCATAGGGGTGTATGGGCTGCAGAGTCAAGCAAGTGTGCATCATTAAGAAAGCGATC 2249
Qy 2224 ACTAAGAAAGATCTGTGTGTGATGTCGAAAGAAAGAAATCTGCGAAGATATCAAGGAT 2283
Db 2250 ACCAAAAAGATCTAAGTGTGAGCGCCAAAGAAAGAAATCTGCGAAGATTAAGAGGAC 2309
Qy 2284 GTAAGAGAGATGAGACGTATGATGTGTGCTGAGATGTCGATTCAGTCTTCAAT 2343
Db 2310 GTCAAGAAATGAAGAGGCTGACGTCAATGCTCAAGATGTGACTGTGCTCTTGAT 2389
Qy 2344 GGGGTTAAGCAACCCGTTTAACTCTGTACATGTAGAGCATTTGCTGCTGCATGCAAGG 2403
Db 2370 GGATGCAAAACACCCGTAAGAACCTGTATTTGACGAAGCTTTTGTCTGTCAAGGT 2429
Qy 2404 ACGTCTGCGCATGATTTGCCATGCTCAACCTTAAGAAAGTGTATGTGTGGGGACCA 2463
Db 2430 ACTCTCAAGAGCGCTCATATGACATTAAGACCTTAAGAAAGGCAAGTGTCTGTGGGGATCC 2489
Qy 2464 AAAAATGCGGCTCTTTAATCATGATGTGCTGAAGATCAATTTTAACATGACATATG 2523
Db 2490 AAAAGTGTGCTTTTAAATGATGTGCTGAAGTGTGCTGAAGTGTGCTTTTAAACAGAGTTTGC 2549
Qy 2524 ACTGAAGTGTACATTAAGAAAGCATCTAAGAGTGTGACACAGATGTAAACCGCATGTC 2583
Db 2550 ACAAGATCTTCCAAAGAAAGCATCTGTGCGGTTGACATTAATGTGTGACTTCCGCTGTC 2609
Qy 2584 TCCACGCTCTTCTACGACAGAGCAATGAAGCGTTAAACCATGTGCTGATTAATATC 2643
Db 2610 TCAACCTGTGTTTACGACAAAGAAAGTGAAGCAAGATTCGAAGAGACATTAATGTG 2669
Qy 2644 ATAGATTCACAGGAGACACAAAGCCGCAAGAGTGTGATTTCTAATCTGTTTCA 2703
Db 2670 ATTACACTACCGGACGATACCAACCTAAGAGAGCATCTCATTTCTCATTTGTTTCA 2729
Qy 2704 GGATGGGTGAACAGCTACAGATTGACTCAAAAATCAGAAATCATGATGCGGCTGCA 2763
Db 2730 GGGTGGGTGAAGCAAGTTCAAATGATTAAGAAAGCAAGAAATATGACGGCAGCTGCC 2789
Qy 2764 TCGCAAGGACTTAAGCGGAAAGCGTTTATGCTGTCAAGTCAAAAGTCAAGAGATCA 2823
Db 2790 TCTCAAGGCTGAGCCCGTAAAGGTGTGTATGCTGTGCTCAAGGTGAAGAAATCTCT 2849
Qy 2824 CTCTACTCGAGACTTCTGAGACAGTGAAGCTGTATCTTACACGACAGAAACGCAAT 2883
Db 2850 CTGTACGACCCACCTCAAGACATGTGAAGCTCTTACGACCCGACGAGAGACGCAATC 2909
Qy 2884 GTCTGAAGAGCTAGCTGTGATCCCTGATTAAGACACTTAAGCTTAATATCCGCG 2943
Db 2910 GTGTGAAGAAACATAGCCGCGGACCCATGATTAAGAAACATGATGCGCAATACCTGTGG 2969
Qy 2944 GATTTCACGCTTCAATGTGACGATGCGAGCGGCAACGAGCCCATTTATGTGACGCTT 3003
Db 2970 AATTTCCTGCGACGATAGAGAGTGTGCAACAGAGCATGTATGATAGGCACATC 3029
Qy 3004 CTGTATTAAGCCGACAGACGCTGATGTGTTCAGAAATTAAGTGAAGCTGTGTGGCGAG 3063
Db 3030 TTGAGAGACCGGACCTTACGAGAGTCTTCAAGATTAAGGCAAAAGTGTGTGTGGCCAG 3089
Qy 3064 GCTTAAAGCCAGTCTTGTGGCCAGCCAAATTTGTGTGACGAGACAGCAAGTGTGGAGACG 3123

Db 3090 GCTTAAAGCCGCTGTGTAAGACCGTGTGCATAGACATGACCACTGAACATGAGCACT 3149
Qy 3124 TTGCACCATTTCAAGCATGACAGAGCTACTCACTGAAATGACATGAACTTTTGC 3183
Db 3150 GTGATTTATTTTGAACGACAAAGCTCACTACAGCAGATATGATTTGAACCAATATGC 3209
Qy 3184 ACCAGTTCTTTTGAAGTGAAGCTGACAGTGGTTATTTTCCGCTCTACCGTGCATT 3243
Db 3210 GTGAGTTCTTTTGAAGCTGATGAGATCTCGGCTATTTTCTGACCCACATGTTCCGTTA 3269
Qy 3244 ACTTAAGGATCAGCATGCGGATTAATCCGCAAGGAAAGACATGTATGGGCTTAATGA 3303
Db 3270 TCCATTGAGATTAATCACTGGGATTAATCCCGCTGTCAATGTACGGGCTGAATAAA 3329
Qy 3304 GAGGTACCAAGAGTGTGACCGGCGATTCGATGATCAAAAGCGTTGACAGAGGC 3363
Db 3330 GAAAGTGTCCGTACGCTCTCTCGAGGTACCCACAACTGCTGGGAGTTGCCATGGA 3389
Qy 3364 AGGTAGCTGATTAAGAAATTAATACCATCAAGACTACTCTCCAAATTAATGTGTT 3423
Db 3390 AGAGTCTATGACATGAACACTGGTACACTGGCAATTAATATCCGCGCATTAACCTAGTA 3449
Qy 3424 CCATTAATGCGCGTTGCCCACTGTTATGTGACCAAGAGACAGGTTCAACT 3483
Db 3450 CTTGTAAACAGAAAGACTGCTCATGTGTAGTCTCCACCATATAGAACCCACAGAGT 3509
Qy 3484 GATCACAGCGATTTCCATATCATAGTAAGAGGCAATCTGTGTGTGATTCGGGATCCT 3543
Db 3510 GACTTTTCTTCAATTCGACCAATTAAGAGGCAAGACTGTCTGTGTGTGGGAAAG 3569
Qy 3544 ATCAGCATTCAGGAAAGAAAGTGAAGTCCATGAGTCCATTCACATTAATACCATAGG 3603
Db 3570 TTGTCCCTCCAGCAAAATGTGTGACTGTGTGTGACACCGGCTGAGGCTACTTCA 3629
Qy 3604 TGTGATCTGATTTGGAATACCTAGCAGTGTGTGATTAATATGACATTAATCTTTGCAAT 3663
Db 3630 GCTGCGCTGATTAAGGACATCCAGGATGTGTGCCAAATATGACATTAATTTGTAT 3689
Qy 3664 GTTAAGACCCCGTACAGAAACATCACTACCAAGAGGAGATCAGGATACACAC 3723
Db 3690 GTGAGAGCCCATTAATTAATCACTATACAGATGATGAAGCAATGCAATTAATACCTT 3749
Qy 3724 AGCATGCTAAGCTGTGAAGCTGTGTCACACCTGAAACACTGGCGAATGATGTGGCTATA 3783
Db 3750 AGCATGTGAACAAAGAGCTGTGTGCTGATGATCCGGCGGAACCTGTGTGACGATA 3809
Qy 3784 GGGTATGCGCTTGTGATTCGCGCAACCGAATATATCACTGTGGGTGACGCTCATTT 3843
Db 3810 GGTATGTGTACGCTGACAGGGCCAGGAAAGCATCAATGTGTATAGCGCGCACTTC 3869
Qy 3844 AGTTTACCCGTGTCTGACGCTTAAGAACACTGCGGAAATCTGAGGTCTCTTGTGTG 3903
Db 3870 AAGTTTCCCGGTATGCAACCGAAATCTCACTTAAGAGAGCGAAGTTCGTTGTGA 3929
Qy 3904 TTCTTCCGCAAGGACCAAGCAACACATGACAGAGACAGACTCGGTGTAGTCTT 3963
Db 3930 TTCAATGGATGACATGCCAGGCGGCTGACCAATCTTATCAAGCTTTATCAACCTTG 3989
Qy 4024 ATCAGAGTGAACATTAAGAGAGCGCTGACCAAGCTATCTTAAATGTGTGTAATAGAAA 4083
Db 4050 GTGTGAAGGATATTTGCAAGGCGCACGGAAGGTGATTAATATGTGTGTAACAGAAA 4109
Qy 4084 GGTCAACAGATTCGAGAGTGTGCGGTGACATGTAACGAAATATGCGCGCTGCTTTGAT 4143
Db 4110 GGAACACTTGGCGAGGAGGTGTGCGAGCGCTGTATTAAGAAATTCGCGAAGAGCTTGAT 4169
Qy 4144 AGACAGCAATAGCTGTGTGGAGCGGTAGAATTGTGAAGACGAAAC--GCTCATATA 4200
Db 4170 TTACAGCGATTCGAATGAGAAAGCGGAGCTGTCAAAAGTGTGACGCTTAACATATCAT 4229

4201 CATGCTGTAGAGACCCAAATTTTCTAAGATGCCGGAACCGAGGCGACCTTAAGCTCCCA 4260
4230 CATGCCGTAGAGCCAAATCTTCAACAAGTTTCGAGGTTGAAGTGAACAACAGTTGGCA 4289
4261 GCTGCTCAATGAGCATAGCTTCATCGTCAACGCTGAGCGGATTAACAATAATACGTA 4320
4290 GAGGCTTATGAGTCTGCTAAGATGTGTCAACGATTAACAATTAACATGCTAGTCCGATT 4349
4321 CCGCTACTGTCAACCGGATCATATCTGTGGCAAGATCCAGTATCATCATATGAT 4380
4350 CCAGTGTGTTCACCGGATCTTTTCCGGAAACAAGATCGACTAACCCAAATCATTTGAC 4409
4381 CACCTGTCTACTGCTTTTGCACATCGATGCGGATGTCACCATATATTTGTTGATTA 4440
4410 CATTTGTCAGACGCTTTAGACACCACTGATGAGTATGACCATATATCTGACGAGACA 4469
4441 CAATGGAGACCGAGATATGAGGCGCATTCACCGCAAAAGAAAGCGTGAATTTCTGA- 4499
4470 AAATGGGAAATGACTCTCAAGGACAGTGGCTAGAGAAAGACATGAGAGATATGTC 4529
4500 -----TGATGACAGCCAGTATGACATTTGCTGTAGGGTCCACCCAAAC 4545
4530 ATATCCGACGACTCTTCACTGACAGACCTGATGACAGCTGTGATAGGCTGATCCGAA 4589
4546 AGCTTTTGGAGAGACAGGTTACTCCGTCATATGAGGCGCAAGTTGATTCATACCTG 4605
4590 AGTTCTTTGGCTGAGAGAGAGGCTACACACACGATGCGAAACTTCTCATATTTG 4649
4606 GAAGGTACACGATTCATCAACGCGCAAGACATTCGCGAAATCATGCAATGTGGCC 4665
4650 GAAGGACCAAGTTTCAACGCGCGCAAGATATAGCAAAATTAATGCCATGTGGCC 4709
4666 AACAAATGTGAGGCTAATGAGAGATTTGCTGTATCATCTGGGGAAGATATGTCAGC 4725
4710 GTTGCAAGGAGGCGCAATGAGAGATGATGATATCTCTGGGAAAGCATGAGCAGT 4769
4726 ATCCGCTCAATATGCCAGTAGAGAGTCAAGAGCGTGTCTCCACTCAACACATTTCA 4785
4770 ATTAGTGTGAATATGCGCGTGAAGAGTGGAGAGCTCTCAACACTAGACGCTGCT 4829
4786 TGCTGTGTATTAATGATGATGAGCGGTGAGCGGTATACAGTTGCGCTTGGCAAGAA 4845
4830 TGCTGTGTATTAATGATGATGAGCGGTGAGCGGTATACAGTTGCGCTTGGCAAGAA 4889
4846 GAACAGTTCGCGGTATGCTATCATCTCTGTGCGGAAGTACAGATACAGCGCTGAG 4905
4890 GAACAAATTAATGCTGTGTCTATCTCTTCCATTTGCGGAAGTATGATCATGCTGTGAG 4949
4906 AAGCTACAGTCAAGCAACAGTCTGTTTTCAGGCGTGTGACCAACGCGCTGTACAGCC 4965
4950 AAGATTCATATGCTCCCAAGCTATATTTGTCTACCGAAAGTCTGCGTATATTCATCA 5009
4966 AGGAAGTA 4973
5010 AGGAAGTA 5017

RESULT 4
US-09-991-258-17
; Sequence 17, Application US/09991258
; Patent No. US20020141975A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Robert
; APPLICANT: Keith, Paula
; APPLICANT: Dryga, Sergey
; APPLICANT: Caley, Ian
; APPLICANT: Maughan, Maureen
; APPLICANT: Johnston, Robert
; APPLICANT: Davis, Nancy
; APPLICANT: Swanson, Ronald
; TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOES WITH MODIFIED HIV GENES FOR USE
; TITLE OF INVENTION: VACCINES

FILE REFERENCE: 0113.000103
; CURRENT APPLICATION NUMBER: US/09/991.258
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 09/902,537
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,995
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 17
; LENGTH: 13584
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. US20020141975A1e =
US-09-991-258-17
Query Match 42.3%; Score 2113.2; DB 10; Length 13584;
Best Local Similarity 64.5%; Pred. No. 0;
Matches 3216; Conservative 0; Mismatches 1748; Indels 24; Gaps 3;
10 ACTAATGATCCATATGGAAGAATTACGTTGATGATGCTGACAGCCGCTATGTC 69
30 AATTACCTACCCAAATGAGAAAGTTCACTTGACATCGAGAAAGACGCCATTCCTC 89
70 AAGTGTACAGCGGAGCTTCCCAATTGATGATCGAAGAGAGGAGGCTGATGACAT 129
90 AGAGCTTGCAGCGAGCTTCCGAGTTGAGGATGAGAGCAAGCAGTCACTGATAT 149
130 GACCATGCAATGCGAGAGCGTTTTCGATGTGCAACAAAGCTCATTTAGAGAGCAAGTC 189
150 GACCATGATATGCAAGAGCGTTTTCGATGTGCAACAAAGCTCATTTAGAGAGCAAGTC 209
190 GACCGGACCAAGTATTTGATGATGAGAGTGGCGCGCTGACATGACATTCAT 249
210 GACCATGCGACAGATCTTGAATGAGAGTGGCGCGCGCGCAAGTATTTCTAAG 269
250 CACCGTATCATTTATGCTGCTATGATGATGAGAGTGGCGCGCGCGCAAGTATTTCTAAG 309
270 CACAAATCATTTATGCTGCTATGATGAGAGTGGCGCGCGCGCGCAAGTATTTCTAAG 329
310 TATGCAAGAACTTAAGAAAGT-----GACATTAACCGCAAGAACTATGAGCTCTAAG 363
330 TATGCAAGAACTTAAGAAAGT-----GACATTAACCGCAAGAACTATGAGCTCTAAG 389
364 GCGGACGCTGCTGGAAGTATGATCAACACGAGAGAGTCTCTGCTGATG 423
390 ATGAAGGAGCTGCGCGCGCTGATGAGCGACCTGACCTGAACTGAGATATGCTGCTC 449
424 CACACAGAGCGGAGTGTAGTATGAGAGTGTAGAGTATACCAAGATGTGAGCGCA 483
450 CACAGAGAGGAGTGTGCTGCTGAGAGGAGGAGTGTGTTTACCAAGATGTGAGCGG 509
484 GTCCATGACCGAGATCATATCAACACGAGCGCTTAAGAGTTAGAGCAATTTACTG 543
510 GTTACGAGACCGAGAGTCTATCAACCAACCAATTAAGAGTTAGAGTGTGCTACTG 569
544 ATAGGCTTTGACACGACCCCTTTATGTATGACAAAACATGAGAGTTCTTACCTTAC 603
570 ATAGGCTTTGACACGACCCCTTTATGTATGTAAGAACTTGAGAGCTATTCATCAT 629
604 AACAGAACTGGGCTGAGAGAGTATGGAACAGCTTACATTTGGCTGGTATCTCA 663
630 TCTACCAACTGGGCGGAGAAACGCTGTAAACGCTGTAAACATAGGCTTATGAGCTCT 689
664 GATCTTACGAGAGCAGGCTTGAAGAACTCTCAATCTCTTGAAGAGAGGCTCCAACT 723
690 GACGTTATGAGCGGCTGACAGAGAGTGTCTTCTTAAAGAAAGTATTTGAAGCA 749
724 ACTAATGATCATATTTCTGCTGTTGTTCAACATTTACAGAGATGATGATCACTGTA 783
750 TCCAACATGTTCTATCTCTGTTGGCTGACCATTAACGAGAAAGAGGAGTACTG 809

QY 784 CGTAGCTGGCATCTTCCAAAGCTGTCCACTTGAAGAAGTCTAATTCAAGTATGA 843
DB 810 AGGAGCTGGACCTGCGCTGTGATTTACTTACGTGGACGAAAATTAACATGTCGG 869
QY 844 TGTGGGACCATTTGTAGCTGTGAAGGGTACGTCACTCAAAAAGATTAACGATCAGCCAGGA 903
DB 870 TGTGAGCTAATAGTTAGTTGGGACGGGTACGTCTGTAAAGATTAAGTATCAGTCCAGGC 929
QY 904 CTATACGCTAAAGTTGAGAACTTGGCGTCCCAATGCAATGCCAGGAGGTTTCTTGAATTGC 963
DB 930 CTGTATGGGAAGCCTTCAGGCTATGCTGCTACGATGACCCCGAGGGATTCCTTGCTGTCG 989
QY 964 AAAGTCCAGATGAGCTGCGCGGAGAGGGTTTCTTTGCTGTGTGATGATGATCA 1023
DB 990 AAAGTACAGACATTCATCAAGGGAGAGGGTCTTTTTTCCCGTGTGACGATATGTCCA 1049
QY 1024 GCCACACTTTGCGATTCAGATGACAGGAAATTCGSCAACTGATAGTGTGATGATGACGA 1083
DB 1050 GCTACATTGTGTGACCAATGACTGGCATCTGGCAACAGATGTCAAGTGGGAGAGCGCG 1109
QY 1084 CAAAACTATTGTTGGGCTCAACCAAGATTTGTCTCAATGTTGAGACCCAAAGAAAT 1143
DB 1110 CAAAACTGCTGTGTGGCTCAACAGGCTATAGTCTCAACGCTCGCACCCAGAGAAAC 1169
QY 1144 ACTAACCAATGCAACTATCTATTATACAGTGGTCCGCGGGGTTTCCAGGTGGGCG 1203
DB 1170 ACCATATACATTAATAATTTACTTTTGCCGATGGCCAGGCAATTTGCTAGGTGGCA 1229
QY 1204 CGTGAACATCTGTGCGCATTTGGAAGCAGAGAAAGAACTAGGGGTGCGGAGCGCACTT 1263
DB 1230 AAGGAATTAAGGAAGATCAAGAAAGTGAAGAGCCACTAGAGACTAGAGATAGACGTTA 1289
QY 1264 ACTATGGGCTCTGCTGGGCTTTCAAGCCAGAAATCAATCATCTTAACAAGGCT 1323
DB 1290 GTCATGGGGTGTGTGGGCTTTTGAAGGCAAGATATACATCTATTATTAAGCGCCG 1349
QY 1324 GGTACGCAAAATTAAGAAAGTACCTGCGCTTTGATCTATTGATGATTCAGGCTT 1383
DB 1350 GATACCCAAACCATATCAAGTGAACAGGCAATTTCCACTATGTGCTCCAGGATA 1409
QY 1384 ACCAGCAGCGGCTCGATATGGGCTTCCCGCTAGAGCTCAAGCTGCTGTGAACCACT 1443
DB 1410 GGCAAGTAACATATGGAGATCGGGCTGAGAACAAAGAAATGTTAGAGAGAGAC 1469
QY 1444 GTCAAAATCCGACCGGCTATTATCAATGGCCGATGTGAGCATCTGCGTGTTCACAGCA 1503
DB 1470 AAGGAGCGTCACTCTCATTTACCGCCGAGGACGTACAAAGGCTAAGTGGCAGCCGAT 1529
QY 1504 GAAGCTGAAGAAAGTGGCTGACGCGGAAGATCAAGAGGCTGCGCAACCTTGCTCCCT 1563
DB 1530 GAGCGTGAAGAGGTGCTGAAGCCGAGAGATTTGCCGAGCTTACACCTTTGGCAGCT 1589
QY 1564 GAAATAGAAAAAGAACCGTAGAGGACAGAACTGATCATTAATGCAAGGACAGAGCA 1623
DB 1590 GATGTGAGAGCCACTCTGGAAGCGATGTGCACTTGATTTAACAAGGCTGGGCC 1649
QY 1624 GGTAGCGTGAAGACCAACGAGACATCAAGGGTGAACATTTACCAAGGCGGAAGAG 1683
DB 1650 GGCTCAGTGAAGACCTCGTGGCTTGAATTAAGGTTACCAAGCTAGATGGGAGAGACAG 1709
QY 1684 ATTGGGTCTTACGGCTATCTTTCAACCCGAGCGGATTTGAATAGTGAAGAAATGCGCGT 1743
DB 1710 ATCGCTTTTACGGCTGTGCTTTCTCCGAGGCTGTACTCAAGAGTGAATTAATCTTGC 1769
QY 1744 ATCCACCATTTGGGGAACAAGTACTGTATATGACTCAAAAGTATGGGCGAGAGATAC 1803
DB 1770 ATCCACCTCTCGCTGAACAAGTCAATAGTATACACCTCGCGCAAAAAGGGGTTAT 1829
QY 1804 AAAGTCAAGCATACACGCTAAGGTCAATTTGACCAAGGAGAGCGGCTCTGTTCAA 1863
DB 1830 GCCGTGAACCATACATGATTAAGTGTGTGACAGAGGACATATGCAATACCGTCCAG 1889

QY 1864 GACTTCCAGGATTTAGTGAAGGCTACGATCGTTTCAACGAGAGGAGATTGTAAC 1923
DB 1890 GACTTTAAAGCTCTGAGTGAAGGTGCCACCATTTGTATCAACGAGCTGAGTTCTTAAC 1949
QY 1924 AGATACCTGACCAACATCGCAATCAACGAGAGCGCTTAAACATGACGAAGTACTAT 1983
DB 1950 AGGTACTGACCACTATTTGCCACACATGAGAGAGCGCTGAACACTGTATGAAGATATTAC 2009
QY 1984 AAGACTTAAGACTCAGACACAGACTCAGATTAAGTCTTCCATATTTAGACGACGAAG 2043
DB 2010 AAAACTGTCAAGCCACAGCAGACAGAGGGAATACCTGTACACATCGACAGAAACAG 2069
QY 2044 TGTGTAAAGCAAGAGACGAGGCTCCTTGCTTAACCGGTATGTAGATTCACCA 2103
DB 2070 TGCCTCAAGAAAGAACTATGCTACTGGGCTAGGCTCAGCGAGCTGTGTGATCTCC 2129
QY 2104 TTTCAAGATTTGCTGACGAGGCTTCAGACAGACGACGACGACCTCAAAATCCCA 2163
DB 2130 TTCCATGAATTTGCTGCTAGAGTCTGAAGAACAGACGACGCTCCTTACCAAGTACCA 2189
QY 2164 ACCATCGGATCTATGAGAGTCCAGGTTCAAGTAAATCTGGAATCATCAAAAGGCTGTG 2223
DB 2190 ACCATAGGGGTATATGGCGTGCAGAGATCAGGAAATCTGGCATTTAAAGGCGAGTC 2249
QY 2224 ACTAAGAAAGATCTGTTGTGAGTGCAGAAAGAAATCGCCAGAAATCATCAGGAT 2283
DB 2250 ACCAAAAAATCTAGTGTGAGCGCCAGAAAGAAATCTGTGCAAAATTTATAGGAGAC 2309
QY 2284 GTAAGAGATGAGACGTATGATGTTGTCTGTAAGACTGTGATTCAGTCTTCTAAAT 2343
DB 2310 GTCAAGAAATGAAAGGGCTGAGAGTCAATGCCAGAACTGTGACGTGCTTGAAT 2369
QY 2344 GGGGTTAAGACCCCGTTAACCTCTGTACATTTGATGAGGCACTTGGCTGCCATGACGG 2403
DB 2370 GGATGCAAAACCCCGTAGACCTGTATATTTGACAACTTTTGTGTATGACAGGT 2429
QY 2404 ACGCTGTGCACTGATTTGCTCATCTGCAACTAAGAAATGTGATTTGTGCGGAGCCCA 2463
DB 2430 ACTCTAGAGCGCTCATAGCTATTAAGACCTTAAAGGCAAGTCTCTCGGGGATCCC 2489
QY 2464 AAACATGCGGCTTCTTTAAATGATGTGCTGAAAGTATCAATTTTAAATGACATATGC 2523
DB 2490 AAACAGTGGCTTTTAAATGATGTGCTGAAAGTATGATTTTAAACAGAGATTTGC 2549
QY 2524 ACTGAATGTACCAATAAAGCATCTCTAGAGAGTGCACACAGCTGTAAACCGCATGCTC 2583
DB 2550 AACAGTCTTCCACAAAGACATCTCGCGGTGCACTAAATCTGTGACTTGGTGTGTC 2609
QY 2584 TCCAGCTCTTCTACGACAGCAATGAGAGCGTTAACCATGTGCTGATTAATAATCATC 2643
DB 2610 TCAATCTTGTTTTACGACAAATAATGAGAACGACGATCCGAAGAGACTTAAGATTGTG 2669
QY 2644 ATAGATTAACAAGGAGACCAAAAGCCGACAAAGATGATTTGATTTAATCTGTTTACA 2703
DB 2670 ATTGACACTACCGGCAATACCAAACTTAACAGAGACGATCTCATCTTGTGTTACA 2729
QY 2704 GGAATGGTGAACAGCTACGATTTGATCAAAAATATCAAAATCAAGATCGGCGCTGCA 2763
DB 2730 GGGGTGTGAAGCACTTGAATTAAGATTAAGGCAACGAAATTAAGAGCGGAGCTGCC 2789
QY 2764 TCCGAAGGACTTACGCGGAAAGGCTTATGCTGTCAAGTAAAGTCAAGCAAGATCCA 2823
DB 2790 TCTCAAGGCTGACCGCTAAGGTGTATGCTGTTGCTGATCAAGGTAAGTAATCTCT 2849
QY 2824 CTCTACTCGAGACTTGTGACAGCTGAACGTGTACTTTACAGCAAGAAAAAGCATT 2883
DB 2850 CTGTACCAACCACTCAGAAACATGTGAACGTCTTACTGACCGGACGAGAGCCGATC 2909
QY 2884 GTCTGAAGACGCTAGTGTGATCTCTGATTAAGCACTTTACAGTAAATATCCCGGG 2943
DB 2910 GTGTGAAAAACATGACCGGAGCCCAATGATTAATAACATGTGCAAGTACCCCTGGG 2969
QY 2944 GATTTCAGGCTTCATTGAGAGACTGACGCGGAAACAGACGCGCATTTATGACGCGTT 3003

Db 2970 AATTTCACCTGCGACAGTATAGAGAGTGGCAAGCAGAGCATGATGCATATGAGGCACTATC 3029
Qy 3004 CTGTGATTAAGCCGCGACAGACAGCTGATGTGTTCCAGATTAAGGTGAAGCTGCTGGCGAG 3063
Db 3030 TTGGAGAGACCGGACCTTACCGACGCTTCCAGAAATAGGCAAAAGGTGTGTTGGCCAG 3089
Qy 3064 GCTTTAGAGCCAGTCTTGCCACCGGCCAATTTGTGCTGACGAGACAGAGTGGAGAGC 3123
Db 3090 GCTTTAGTGGCGGTCTGTAAGACCGCTGGCATATGACATGACCATGAACTAGAACT 3149
Qy 3124 TTGCAACCATTCACAGATGACAGAGGCTACTCACTGAAATGCACTGAACTTTTTC 3183
Db 3150 GTGGATTATTTTGAACCGAACCAACTCACTCAGAGAGATGATTTGTAACCACTATGC 3209
Qy 3184 ACCAGGTTCTTTGAGTAGACCTGGAACAGTGGTTATTTTCGGCTCTAACCGTGCACCT 3243
Db 3210 GTGAGGTTCTTTGACTGATCTGATCTGACCTCGGTCTATTTTCTGCAACCACTGTCCT 3269
Qy 3244 ACTTACAGGGATCAGACCTGGGATTACTGCGCAGGAAAGAACATGTATGGCTTAATGA 3303
Db 3270 TCCATTAGGAATATCATCTGGGATTACTCCCGTGGCTTAACATGTAGCGGCTGAATAA 3329
Qy 3304 GAGGTAGCAAAAGAGTGTACGCGGATATCCGTGCATCAAAAACGGTTGACACAGGC 3363
Db 3330 GAAGGTGTCCTGACCTCTCTCGCAGGTACCCACAACCTGCTCGGGCAGTTGCCACTGGA 3389
Qy 3364 AGGGATGCTGATATTAAGAAATATATCAATCAAGGCTACTCTCCAAACAATTAATGTGTT 3423
Db 3390 AGAGTCTATGACATATACATGTGTACCTGCGCAATTTGATCTCCGCGCAATTAACCTAGTA 3449
Qy 3424 CCATTAAATCGCGGTTGCCCACTGTTGATGTTGACACCAAAAGACAGGGTCAACT 3483
Db 3450 CCGTAAACAGAAAGCTGCTCATGCTTTAGTCTCCACCAATATGAACCCACAGAGT 3509
Qy 3484 GATCAGACGGATCTCTATCTAAGATGAAGGCCAAATCTGTGTGTGATCGCGCATCT 3543
Db 3510 GACTTTCTTCAATCGTCAGCAAAATTTGAAGGCGCAAACTGCTGTGTGCGGAGAAAG 3569
Qy 3544 ATCAGCATTCACAGGAGAAAGTAGATGTCATGAGGTCATTTGCCCACTAATACCTACG 3603
Db 3570 TTGTCCGTCACAGGAAATGTTGATGTTGTGACACCGGCTCGAGGCTACCTTCAGA 3629
Qy 3604 TGTGATCTGATTTTGGGATATCTAGCCATGTGCGTAAATATGACATTATCTTGTCAAT 3663
Db 3630 GCTGGCTGGAATTAAGCATCCAGTGATGTGCCAAATATGACATATATTTGTAAAT 3689
Qy 3664 GTTAGAGCCCGTACAGAAACCATCACTACCAACAGTGCAGATTCACGTAATCCACAC 3723
Db 3690 GTGAGGACCCCATATTAATATACATCATATCAGAGTGAAGACCATGCCATTAACTT 3749
Qy 3724 AGCATGCTAACGCTGAAGGCTGTCCACACCTGAAACATGCGGAAACATGTGTGCTATA 3783
Db 3750 AGCATGTGACCAAGAAAGCTTGTGTGATCTGAATCCCGGCGAACTGTGTGACATA 3809
Qy 3784 GGGTATAGGGCTTGCTGATCGCGCAACCGAATATCATCACTGCGGTGACACGCTCATTT 3843
Db 3810 GGTATATGTTACGTGACAGGCGCAGGAAGCATTTGTGTCTATAGCGCGGAGATTC 3869
Qy 3844 AGGTTTACCCGTGTCTGACGCTTAAGAACATGCGGAAATATGAGGTTCTCTTCGTG 3903
Db 3870 AAGTTTCCCGGGTATGCAAAACCGAATCTCACTTGAAGAGACGGAAGTTCTGTTGTA 3929
Qy 3904 TTCTTCCGCAAGACAAAGGCAACACACATGACAGACAGCTCGGTATGATGCTT 3963
Db 3930 TTCAATTGGGTAGATCGAAGGCCCGGTAGCGCAATCTTTCACACCTTTCACACCTTG 3989
Qy 3964 GACAACATCTATCAAGGCTCAACAGTACAGAGGAGGAGGCTCCAGCGTACAGAGT 4023
Db 3990 ACGAATCTTATACAGGTTCCAGACTCCACGAAGCCGATGTGCACTCATATATATG 4049
Qy 4024 ATCAGAGGTGACATTAGCAAGAGCGCTGACCAAGCTATCGTTATGCTCTAATAGCAA 4083

Db 4050 GTGCGAGGGATATTGCCAGGCCACCGAAGAGTGAATTAATATGCTGCTAACAGCAA 4109
Qy 4084 GGTCAACAGGTTCCGAGTGTGCGGTGCTCACTGTACCGAAATGCGCGGCTGCTTTGAT 4143
Db 4110 GACAACTTGCGAGAGGGGTGTGCGAGCGCTGTATAGAAATTTCCGGAAAGCTTCAT 4169
Qy 4144 AGACAGCCAAATAGCTGTGCGGACGGCTAGACTGTGGAAGACGAAC--GCTCATATA 4200
Db 4170 TTACAGCCGATCGAAGTGAAGAAAGCGGACTGGTCMAAGGTGACGCTAATATCATTT 4229
Qy 4201 CATGCTGAGACCCCAATTTTCTTAAGTGTCCGGAACCGGAGGCGACCTTAAGTCCGA 4260
Db 4230 CATGCCGTAGGACCAAACTTCAACAAAGTTTCGGAGTTGAAGTGACAAACAGTTGCA 4289
Qy 4261 GTCGCTCATGAGATGATGCGGCCATGCTCAACGCTGACGGGATTAACAAAATATCAGTA 4320
Db 4290 GAGCTTATGATGCTATGCTTAAGTTGTCAACGATTAACAATTAACATCAGTACGATTT 4349
Qy 4321 CCGCTACTGTCAACCGGATCTATTTGTTGTTGGGCAAAAGATGAGTATCATTTGAT 4380
Db 4350 CCACTGTTGTCAACCGGATCTTTTCCGGAAACAAAGATGACATAACCATATTTGAAC 4409
Qy 4381 CACTGTTCACTGCTTTGACACTACAGATGCCGATGTCAACATATATTGCTTGATAA 4440
Db 4410 CATTTGCTGACAGCTTTTGAACACACATGATCAGATGAGCCATATCTGACAGGACAA 4469
Qy 4441 CAATGGGAGACAGGATTAATGAGGCCATTACCGGCAAAAGAGCGTGAATTCGTGA- 4499
Db 4470 AAATGGAAATGACTCTTAAGAAAGCAGTGGCTAGGAAGAGCAGTGAAGATATGC 4529
Qy 4500 -----TGATGACAGCAGTAGACATTTGCTGTGAGGGTCCACCAAC 4545
Db 4530 ATATCCAGCATCTTCAAGTACAGAACTGTATGACAGAGCTGTGAGGTCATCCGAAG 4589
Qy 4546 AGCTTTTGGAGGACACAGGTTATCTCCGTCAATGAGGCAAGTTTATTCATCTG 4605
Db 4590 AGTTCTTGGCTGGAAGGAGGAGGCTACAGCAACAGCGATGCAAACTTCTCATATTTG 4649
Qy 4606 GAAGTACAGATTCCATCAGACCGGCCAAGGCAATGCGGAAATCCATGCAATGTGGCC 4665
Db 4650 GAGGAGACAAATTTCACAGGCGGCGCAAGATATGAGAAATTAATGCCATGTGGCC 4709
Qy 4666 AACAAATCGAGGCTAATGACAGATTTGCTGTATCACTCGGGGAGAGTATGTCAGC 4725
Db 4710 GTTGCAACGAGGCGCAATGACAGATATGATATATCTCGGAGAAAGCATAGCACT 4769
Qy 4726 ATCGCTCCAAATCCCAAGTAGAGAGTCAAGGCGCTGCTCACTCACACTTCCA 4785
Db 4770 ATTAGTGAATATCCCGTGAAGAGTGGAAAGCTTCTCACCACTAGCAGCTGCT 4829
Qy 4786 TGCCGTGTATTAACGCTATGACCGCTGAGGCGCTATACAGTTGCGCTTGCAGAA 4845
Db 4830 TGCTGTGATCAATGCTATGCTATGCTCAAGAAAGATACAGGCGCTTAAGGCTTCAGTCA 4889
Qy 4846 GAACAGTTCGCGATGCTCATCTCTGTTGCGAAGTACAGATACAGGCGTGCAG 4905
Db 4890 GAACAAATTAATGTGTGCTCATCTTTCATTCGCGAAGTATAGATATACGTGTGAG 4949
Qy 4906 AAGCTACGTGACAGCAACAGCTCTGTTTCAGGCGCTGTACCAACCGGCTGACACCC 4965
Db 4950 AAGATCCAAATGCTCCAGCCTATATGTTCTCACCGAAAGTGCCTGATATATCATCA 5009
Qy 4966 AGGAAGTA 4973
Db 5010 AGGAAGTA 5017

RESULT 5
US-09-991-258-2
; Sequence 2, Application US/09991258
; Patent No. US20020141975A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Robert

Db 1681 GTGCTTTCGCCAGGCGTGTACTCAAGAGTGAATAATATCTTGATCCACCTCTCGCT 1740
Qy 1759 GAACAGACTGTGTATGATCTACAAAGGTAGGGAGGAGATACAAAGTGAAGCATAC 1818
Db 1741 GAACAGACTGTGTATGATCTACAAAGGTAGGGAGGAGATACAAAGTGAAGCATAC 1800
Qy 1819 CACGGTAAAGTCTATGTATCCAGAAAGGAGCGGCGCTTCTCAAGACTTCCAGGCTTGG 1878
Db 1801 CATGTTAAAGTATAGTGTGCCAGAGGACATGCATACCCGTCAGAGACTTTCAGCTCTG 1860
Qy 1879 AGTGAAGCGCTACATCGTTTCAAGAGAGGAGTTGTAACAGATACCTGACACAC 1938
Db 1861 AGTGAAGCGCTACATCGTTTCAAGAGAGGAGTTGTAACAGATACCTGACACAC 1920
Qy 1939 ATCCGAATCAACGAGAGGAGCGCTTAACATGACAGAGTACTATTAAGTCTAAAGCT 1998
Db 1921 ATTCGACACATGAGGAGCGCTGAACATGATGAAGATATTAACAAACTGTCAAGCC 1980
Qy 1999 CAGGACACAGACTCAGATTAAGTCTTCAATATTAAGAGAGAGGAGTTGTAAGCGAGAA 2058
Db 1981 AGCGAGACAGACGCGGAATCTCTGATCAGATCAGAGAGAAACAGTGTCAAGAAAGAA 2040
Qy 2059 GACGCAAGTCCCTGTGCTTAACCGGTATCTGTAGATCCACATTTACAGATTTGCG 2118
Db 2041 CTAGTCACTGGGCTTACGAGCGAGCTGTGTGATCTCCCTTCATGAAATTCGCC 2100
Qy 2119 TACGAGACTCTCAGACACGACAGACAGACCTCAACAAAGTCCCAACCATGAGACTAT 2178
Db 2101 TACGAGACTCTCAGACACGACAGACAGACCTCTCTTCAACAAAGTCCCAACCATGAGGCTGTAT 2160
Qy 2179 GAGAGTCCAGGTTGAGTAAATCTGGAATCAATCAAAAGCGTGTGACTTAAGAAAGATCTG 2238
Db 2161 GCGGTCCAGAGATCAGGCAAGTGTGATCATTAAGAGAGAGTCAACAAAGAAATCTA 2220
Qy 2239 GTTGTAGTGCAGAAAGAAAACTGCCAGAAATCATCAGGAGTAAAGAGATGAGA 2298
Db 2221 GTGTGAGCGCCAGAAAGAAAACTGTGCAAGAAATTAAGGAGAGTCAAGAAATGAAA 2280
Qy 2299 CGTATGATGTGTGCTGTAGAGCTGTGATTCAGTGTCTTCAATAGGGGTTAAGACCCC 2358
Db 2281 GGGGTGAGCGCTCAAGAACTGTGATCTCAGTGTCTTGAAGATGAGAAACACCCC 2340
Qy 2359 GTTAACTCTGTATCATGTATGAGGATTTGCTGCTGCATGACAGGAGCGTGTGCACTG 2418
Db 2341 GTAGAGACCTGTATATGACGAGAGCTTTGCTTGTCAAGGATCTTCAAGACGCTC 2400
Qy 2419 ATTGTCATCTCAAACTTAAGAAAGTGTATTTGTCGGGGAGCCCAAACTGCGGCTTC 2478
Db 2401 ATAGCATTAATAGACTTAAGAAAGGAGTGTCTGTGGGGATCCCAACAGTGGCTTTT 2460
Qy 2479 TTTTAACTGATGTGCTGTAAGATATTTTAACTCATGACATATGACATGAAGTGTACAT 2538
Db 2461 TTTTAACTGATGTGCTGTAAGATATTTTAACTCATGACATATGACATGAAGTGTACAT 2520
Qy 2539 AAAAGCATCTCTAGAGGTGACACAGACTGTAAACCGCATGTCTCAACCTCTTCTAC 2598
Db 2531 AAAAGCATCTCTCGCGTTGCACTAAATCTGTGATTTGGTGTCTCAACCTTGTTCAC 2580
Qy 2599 GACAGAGCATGAAAGCGGTTTAACTCATGTGCTGTAATAATCATATGATATCCACAGG 2658
Db 2581 GACAGAGCATGAAAGCGGTTTAACTCATGTGCTGTAATAATCATATGATATCCACAGG 2640
Qy 2659 AACCAAAAGCGGCAAAAGATGATCTGATTTCACTGTCTTCAAGAGATGGGTGAACAG 2718
Db 2641 AGTACCAAAAGCGGCAAAAGATGATCTGATTTCACTGTCTTCAAGAGATGGGTGAACAG 2700
Qy 2719 CTACAGATTTGACTCAAAATCAAGAAATCAATGATGGGCTGCAATCCCAAGAGATTACG 2778
Db 2701 TTGCAAAATGATTAAGAAAGCAAGAAATTAAGAGGAGCGCTCTCAAGAGGCTGAC 2760
Qy 2779 CGGAAAGCGGTTTATGTGTGCTGATGATCAAGAAATCCACTTACTGCGAGACT 2838

Db 2761 GGTAAAGTGTGTATGCGGTTCCGTTACAGAGTGAATGAATAATCTCTGTACACCCACC 2820
Qy 2839 TCTGAGCAGCGTAAGCTCTTACTTACACGACAGAAAAAGCATTTGTGTGAAGCGCTA 2898
Db 2821 TCAGAACTATGAAAGTCTTACTTACACCGACAGAGAGCGCATGTGTGTGAACACTA 2880
Qy 2899 GCTGTGTATCCCTGGATTAAGACACTTACACTTAAATATCCCGGGATTTTACAGGCTTCA 2958
Db 2881 GCCCGACACCATGATTAAGAAACACTGATCTCCAAAGTACCTGGGAATTTTCACTGCCACG 2940
Qy 2959 TTGACAGCTGACGCGGAAACAGACGCCATTAATGACGCGCTTTGATTAAGCGGAG 3018
Db 2941 ATAGAGAGTGGCAAGCAGACATGATGATCCATATGAGGACATCTTGAAGAGACCGAC 3000
Qy 3019 ACAGCTATGTGTTCCAGAAATTAAGTGAACCTGTGCGGGGAGGCTTTAGAGCAATC 3078
Db 3001 CTACCGAGCTCTTCCAGAAATTAAGTGAACCTGTGTGGGCAAGGCTTTAGTGTGCGGTG 3060
Qy 3079 TTGGCCACGCGCAACATTTGTCTGACGAGACAGAGTGGAGAGCTTGCACCATTTCAAG 3138
Db 3061 CTGAAGACCGCTGGCATGATGATGACATGACATGAACATGGAACATCTGTGATTTGAA 3120
Qy 3139 CATGACAGAGCTACTCACTGAATGACACTGAATCTTTTGACACAGGTTCTTTGGA 3198
Db 3121 ACAGCAAAAGCTCACTGACAGAGATGATTTGAACCAATATGAGTGTGAGGTTCTTTGGA 3180
Qy 3199 GTAGACTGACAGTGGGTTATTTTCCGCTCTTACCGTGCACATTAATTAAGGATCAG 3258
Db 3181 CTGATCTGAGCTCCGCTATTTTCTGACACCACTGTCTCGTTATCATTAAGAAATAT 3240
Qy 3259 CACTGGGATTACTGCCAGGAGAAACATGATGAGGCTTAATAGAGTGTGACAAAGAG 3318
Db 3241 CACTGGGATTACTGCCAGGAGAAACATGATGAGGCTTAATAGAGTGTGACAAAGAG 3300
Qy 3319 TTGTACGCGCATATCCGTGATCAACAAAGCGTTGACAGAGCAGGATGATATA 3378
Db 3301 CTCTCTGAGGATACCCCAACACTGCTGGGCAAGTTGCACTGGAAGATCTATGACATG 3360
Qy 3379 AGGAATTAATCACTCAAGAGATCTCTTCAACATTAATGTGTCTTAAATGCGCGG 3438
Db 3361 AACACTGTATACCTGCGCAATTAATGATCCGCGATTAACCTGATTAACAGAAAGA 3420
Qy 3439 TTGCCCACTGTTGATCTGTGAACCAAAAGACAGGTTACATCAACAGCGGATTT 3498
Db 3421 CTGCTCATGTGTTATGCTCTCAACATTAATTAACCAACAGAGTACTTTTCTCATTC 3480
Qy 3499 CTATCTAAGTGAAGGCGAAATCTGTGTGTGATTCGCGATCTCATACGATTCAGAGG 3558
Db 3481 GTCAAGAAATTAAGAGGAGAACTGTCTGTGTGTGCGGAAAGTTGTCCGTCCAGGC 3540
Qy 3559 AAGAAATGATGATCAATGAGTCTCATTTGCCCATTAATCATCAGGTGTGATTTG 3618
Db 3541 AAAATGTGTGATCTGTGTGTCAGACCGGCTGAGGCTACTTCAAGCTCGGCTGATTTA 3600
Qy 3619 GAAATACCTGACATGTCGTAATTAATGACATTAATCTTTGCAATGTTAAGAACCCGAC 3678
Db 3601 GGCATCCAGGTGATGTGCCAAATTAATGACATTAATTTGTTAATGTAAGAACCCCATAT 3660
Qy 3679 AGGAACATCACTCAACAGTGCAGAGATCAACCTTACACACAGATCTTAACGTGT 3738
Db 3661 AAATACATCACTTACAGAGTGTGAAGACATGATCAATTAAGTTATTAACCAAG 3720
Qy 3739 AAGCTGTCAACCTGTAACACTGGGGAACATGTGTGCTATAGGATATGGCTTGTCT 3798
Db 3721 AAAGCTTGTGTGATCTGAAATCCCGGCGGAACCTGTGTCAACATAGGTTATGATACCT 3780
Qy 3799 GATGCGGAACCGGAATTAATCATCACTGCGGTGACAGCTCATTAAGTTAACCGGTGC 3858
Db 3781 GACAGGCGCAGGAAGCATATGTGTGCTATTAACGCGGACATTCAGATTTTCCGGGTA 3840
Qy 3859 TGTGAGCTTAAGAACTGTGCGAAATAATGAGGTTCTTCTGTCTTCTGCGAGAGAC 3918
Db 3841 TGCAACCGAAATCTCACTTGAAGAGACGGAATTTCTGTTTGTATTCATTGGGTACGAT 3900

```
QY 3919 AACGGACACACATGACAGAGACTCGGTGAGTTCAGACATCTATCAA 3978
DB 3901 CGCAGGCGCCGTACGACCAATCTTACAGCTTTCATCACTTACCACTTTATACA 3960
QY 3979 GGGTCACCAAGTACGAGGCGAGAGCTTCACGCTACAGAGTATCAGAGTGACATT 4038
DB 3961 GGTTCAGATCCACGAGAGCGGATGTGACCTCATATCATGTGTGCGAGGGATATT 4020
QY 4039 AGCAGAGCGGTGACCAAGCTATGTTATCTGCTTAATAGCAAGGTCAACGAGTTCC 4098
DB 4021 GCCAGCGCCACCGAAGAGTATTAATGCTCTTAACAGCAAGAGCAACCTGGCGGA 4080
QY 4099 GGAAGTGGCGGTGACTGTACCGAAATGCGCGCTGCTTGTAGACAGCAATAGCT 4158
DB 4081 GGGGTGTGGAGCGCTGTATTAAGAGTTCCCGAAGAGCTTCGATTTACCGCCATGAA 4140
QY 4159 GTCGGGACGGCTAGACTTGTGAAACAGAACCC--GCTCATCATACATGCTGTAGACC 4215
DB 4141 GTAGAAAAGCGGAGCTGTCAAAAGGTGACCTAAACATATCATTTGCGGTAGACCA 4200
QY 4216 AATTTTCTAAGATCCGGAACCGGAGGCGACTTAAGCTCGGAGCTGCTACATGAGC 4275
DB 4201 AACTTCAACAAAGTTTCGAGAGTTGAAGTGACAAACAGTTGGCAGAGGCTTATGAGTCC 4260
QY 4276 ATACGTCATCGTCAACGCTGAGGAGATTACAAAATATGATACGCTACGTCACACC 4335
DB 4261 ATGCTTAAGATTTGTCAACGATTAACAATTACAGTACGAGTATTCATCTTTGTCACC 4320
QY 4336 GGCATCTATTCTGTGGCAAGATGAGTGTATGATGATCATTCATCACTGTTGACGCT 4395
DB 4321 GGCATCTTTTCCGGGAACAAAGATCGACTAACCAATCATTTGACATTTGCTCAGCT 4380
QY 4396 TTGCACTAAGGATGCGGATGTCACCATATATTGCTTGTGATTAACAATGGAACCAAG 4455
DB 4381 TTACACACCACTGATGAGATGATGACATATATGACAGGACAAAGAAATGGAAATGACT 4440
QY 4456 ATATCGAGGCACTTCAACCGCAAGAAAGCGTCGAAATTTGGA-----T 4500
DB 4441 CTCAGAGAGAGATGCTAGAGAGAGAGAGAGTGAAGAGATATGATATCCGACGACTCT 4500
QY 4501 GATGACAAAGCCAGTAGACATTTGCTGAGGGTCCACCCAAACAGCTCTTTGGCAGGC 4560
DB 4501 TCAGTGAAGAACTGATGACAGCTGGTGAAGGGTGAATCCGAAGAGTTCTTTGGCTGGA 4560
QY 4561 AGACCAAGTTACTCGCTCAATGAGGCGCAAGTTGATATCTGGAAGGTACAGATTC 4620
DB 4561 AGGAAGGCTACACACAAAGGATGGCAAACTTTCTCATATTTGGAAAGGACCAAGTTT 4620
QY 4621 CATGACACCGCCAAAGACATTTGCGAAATTCATGCAATGTGGCCCAAACTTGAGGCT 4680
DB 4621 CACCAAGCGCCAAAGATATGACAAATTAATGCTATGTGCGCTGGTTCGACAGAGGCC 4680
QY 4681 AATGAGCAGATTTGCTGTATCATCTGTGGGAGAGATATGTCACAGATCCGCTCCAAATGC 4740
DB 4681 AATGAGCAGATGATCATGTATATCTCGGAGAAAGCATGACAGATTAATGATCGAAATGC 4740
QY 4741 CCAATGAGAGATGAGAGGCTGTGCTCACTACACACTTCCATGCTGTGTATATAC 4800
DB 4741 CCCGTCGAAGAGTGGAAAGCTTCCACACACACTGACAGCTGCTGTGTGTGATTCAT 4800
QY 4801 GCTATGACGGCTGAGCGGCTATACAGGTTGCGCTTGGGAAAGAAACAGTGTGCGCTA 4860
DB 4801 GCCATGACTCCAGAAAGAGTACAGCGCTTAAAGCCTTACGCTCCAGAACAATTAATCTGTG 4860
QY 4861 TGCTCATCATCTCTGTGTCGCAAGTACAGATTCACAGGCGTGCAGAACTCAAGTGCAGC 4920
DB 4861 TGCTCATCTTTTCATTTGCTCGGAAGTATGAAATCACTGTGTGCGAAGATCAATATGCTCC 4920
QY 4921 AAACCAAGTCTGTTTTCAGGCGTGTACCAACGCGTGTACACCCGAGAAAGTA 4973
DB 4921 CAGCTATATATTTCTCACCGAAAGTGTGCTGATATATTCATCCAAAGAAAGTA 4973
```

```
RESULT 6
US-09-994-412-3
; Sequence 3, Application US/09994412
; Publication No. US2003059943A1
; GENERAL INFORMATION:
; APPLICANT: Certia, Ulrich
; APPLICANT: Landstrom, Kenneth
; TITLE OF INVENTION: Inhibition of Expression of a Target Gene
; FILE REFERENCE: 20787
; CURRENT APPLICATION NUMBER: US/09/994,412
; CURRENT FILING DATE: 2002-03-12
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 10610
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pSFV2gen(PD) vector
US-09-994-412-3

Query Match      31.3%; Score 1567.2; DB 11; Length 10610;
Best Local Similarity 58.4%; Pred. No. 0;
Matches 2906; Conservative 0; Mismatches 2038; Indels 36; Gaps 8;
```

```
QY 31 AGAATTCAGCTTGAATGATGCTGACAGCGCGTATGTCAAGTCTTACAGCGAGCTT 90
DB 96 AAAGTCATGTGATATTTAGAGCTGACAGCCCATTCATCAAGTCTTTGCAAGAGCATTT 155
QY 91 CCAAAATTTGAGATCGAAGCAAGGAGCTGACATGACATGACCAATGCGAATGCGAGCG 150
DB 156 CCGTGTGAGAGTGAGTATGATGAGCTGACAGCTCAACCAATATCATTGCAATGCGAGCA 215
QY 151 TTTTGCATGTGCAACAAGCTCATTTGAGAGGAACTGACCGGACCAAGTTATTTG 210
DB 216 TTTTGCACCTGGCTACCAATTTGATGACAGAGAGACTGACAAAGCACACATCATCTTG 275
QY 211 GACATTGGAAGTGGCGGCTGACAGATGACATTTCCATGACCGCTTCAATTTGATGTC 270
DB 276 GATTCGCGAGTGGCTTCCAGAGAGATATGTCGACAAATTCACCTGCTATGTC 335
QY 271 CCTATGATTAAGCGCTGAAAGACCGGACAGACTCAACAGTATGACAAAGACT-----T 324
DB 336 CTTATGCGGAGCGGAGAAAGACCCGAAAGGCTGATGCTACGAAAGAAATGCGAGCG 395
QY 325 AAGAAAGTACATTAACCGACAAAGATAGCTCTTAAGCGGCGACCTGTGGAATGC 384
DB 396 GCTCCGGAAGGTGCTGATAGAGATGCGAGAAATTCACCGACTGTGACAGCCGTC 455
QY 385 ATGTCAACACAGACGAGAGAGCTCCATCTCTGTATGACACAGCGCGACGCTGATAG 444
DB 456 ATGGCTAGCGCAGACGCTGAATCTCTTACCTTTGCTGTGATGACGATCAGTGTGCT 515
QY 445 TACTTTGGAAGTGAAGATGATACCAAGATGTGTACAGCTGATGACAGCAGCATATC 504
DB 516 ACGGACCGGAAGTGGCGGTATACAGAGAGTGTATCTGTATATGACCAACATGCTG 575
QY 505 TACCACAGGCGCTTAAAGAGTGAAGCAATTTACTGATAGGCTTTGACAGACCCCT 564
DB 576 TACCATCAGCGATGAAGAGTGTGAAACGCGATGATGATGAGTGTGATGACACACCCG 635
QY 565 TTTATGTACAAAACATGGAAGGTTCTTACCTTACTTACCAACAGACTGGGCTGAGAG 624
DB 636 TTTATGTTTGAACGCGCTAGACGCGGCTATTCACCTTACGCAACAAATGCGCGAGAG 695
QY 625 AGAATATTGAAGACGCTTAACATTTGCGCTGCTGATACCTCAGATCTTTCAGAGAGCGCTT 684
DB 696 CAGGTGTACAGGCGACAGAAATATGAGACTGTGTGACACATCTTTGATGAGGGAAGATC 755
QY 685 GAAAACTCTCATCTTTAGAGAGAGAGGCTCCAACTCTAATTAAGATCATTTCTCG 744
DB 756 GGCAAACTGTTCATTTCTCCGCAAGAGCAATGAAACCTTGCGACACAGTCAATGTTCTCG 815
```

OY		745	GTTGGTTCACAAATCTATACACAGAAATGATCATCGTTAATACGTAGCTGGCACTCTCCAAAC	804
Db		816	GTAGGATCTAACCTTTACTCTGAGCAGAAAAGCTATCTGAGAGCTTGACACTTACCCTCC	875
OY		805	GTGTTCCAATTGAAGAAGTAAGTCTATCTTCACTCACAGGTAGATGTGGACCATTTGTCACTGT	864
Db		876	GTATTCACCTGMAAGTAAACAAATCCTTTACCTTAGGTGGATACCATGSTATCANGT	935
OY		865	GAAGGATACGTCACTAAAAAGATAAGATCAGCCCAAGACATATACGTTAAAGTTGAGAAC	924
Db		936	GAAGGATACGTATTAAGAAAAATCATATGTGCCCGCCCGCTGTATGAGTAAACGGTTAGGG	995
OY		925	TTCGGGTCACCAATGTCATCGCAGAGGGTTTTCTTGAATTTCTTCAAAGTCAACGATACGCTGGCG	984
Db		996	TACGCCGTGACGTATATCACGGCGAGGGATTTCTTAGTGTGCMAAGACACAGACACTGTCAA	1055S
OY		985	GCGCAGAGGGTTTTCTTTTGTGTGTGTACGTATGTACCAAGCCACATTTGGCATGATG	1044
Db		1056	GGAGAAAGATCTCATTTCCCTGTATGACCTAACGTCCTCCCAACCATCTGTGATTAATG	1115
OY		1045	AACAGGATTTCTGGCACTGACGTTTAGTGTGATGATGACGACAAAACCTATTTGGTTGGGCTC	1104
Db		1116	ACTGCGATATACGACCAACGACGTCACACCGAGAGACGACAGAACTTTATATGTAGATTG	1175
OY		1105	AACCAAGGATTTGTCGTAATGTATGTGAGACGCAAAAMAAATTAATCACATACAGATCATAT	1164
Db		1176	AATCAGAGATATAGTTGTGACGGAAGAACACAGCGAACACTAACAGATGAAGACATAT	1235
OY		1165	CTATTATCCAGTGTGTGCCACGAGCGTTTTCCAGTGTGGGCGCTGAACATGTGCCGACTTG	1224
Db		1236	CTGCTTCGATGTGTGGCCGTGCATTTAGCAATGTGGGAGAGGAATACAAAGCAGACCTT	1295
OY		1225	GACGACGAGAAAGAACTAGAGGGGTGGGGAGCGCATCTTACTATGGCTGCTGGGCT	1284
Db		1286	GATGATGAAAAACCTCTGGGTGTCCGAGAGAGTCACTTACTTGTGCTGCTTTGTGGCA	1355S
OY		1285	TTCAAGACCCAGAAAAATCAATCCATCTAACAGAAAGCTGTGACGCAAAATTAAGAA	1344
Db		1356	TTTAAAAAGAGGAAGATGACACACATGTACAAAGAACGACACCCMAACAAATAGTGAAG	1415
OY		1345	GTACCTGCGGCTTTGATCMCAATTTGTGATTCAGCGCTTACCAAGCAAGGGCTCGATATG	1404
Db		1416	GTGCTTTAGAGTTTAACTTCGTTGTCATCCGAGACCTTAGTGTCTACAGGCGCTCGCAATC	1475
OY		1405	GGCTTCCGCGTAAAGCTCAAGCTGTGCTTGAACCACTGTCAAAACCGCACCGGCTAT	1464
Db		1476	COAGTCAGATCACCGATTAAGATGCTTTTGGCAGAAAGACCAAGC---GAGAGTAAATA	1532
OY		1465	ACAAATGCGCGATGTGAGCAATCTGCGTGGCTTACAGCAAGAACTGAAGAAAGTGGCTGCA	1524
Db		1533	CCTGTTCGACGCGCTGTCAAGCCAGGGATGTGAACAAAGAGAGAAAGAGATTGGAG	1592
OY		1525	GCAGAAAGATACAGAGAAAGCCCTGCGACCCCTTGCTCCCTGAATAATGAAAAAGAGCCGTA	1584
Db		1593	GCCGAGCTGACTAGAGAAAGCTTACACACCTCTGCTCCCATTCGCGCGCGGAGAGCGGA	1652
OY		1585	GAGCGAGAATGAGACTCATTAATGCAAG------GCAGAGCAGGATAGCGTGGAGCA	1638
Db		1653	GTCCGTGACGTCGACGTTGAAGAACTAAGATACGACGAGGAGCGGGGTCGTGGAAACA	1712
OY		1639	CCAGAGGACACATCAGAGGTGACAAAGTTAACCAAGCCGAAGAAAGATTGGGTCTTAAGCT	1698
Db		1713	CCTGCGACGCGGTTGAAGATCACCGCACAGCCGAACGACGTACTATAGGAAATTAAGTA	1772
OY		1699	ATACTTTACCCCAGCGGCTATTTGAATATGTGAAAAAACTGGGTGTATCACCCATTGGCG	1758
Db		1773	GTTCTGTCCCCGACAAACCGTGTCTAAGAGCTCAAGATTGGCCCCCGTGCACCTCTAGCA	1832
OY		1759	GACAAAGTACTGTAAATGATCAACAAGAGTAGGGACGAGAGATCAAAAGTGAAGCATATC	1818
Db		1833	GAGCAGGTGAAAAATATATAACATATACGAGAGGGCGCGCGGTATACAGAGTGAACGAGAT	1892

QY	1819	CACGGTAAAGGCTATTGTACACGAAGGGACGGCGGGTCCCTGTGTTCAAGCTTCACAGCAATG	1878
Db	1893	GACGGCAGGGGTCTACTACATCAATGTGGATCGGCATTCCTGGGTCCCTGAGTTTCAGGCTTTG	1952
QY	1879	AGTGAAGCGCTACGATCGTTTTCACAGAGGGAGTTCGTAAACAGATACCTGCACAC	1938
Db	1953	AGCGAGAGCGGCATTAATGTGTATACAGAAAGGAGTTTCGTCAACAGAAACTATACAT	2012
QY	1939	ATCGCAATCAACGAGGAGCGCTTAACCTGACGAGAATGCTATTAAGACTGTAAAGACT	1998
Db	2013	ATTGCGCTTACGGAGACCTCGCTGAACACCGACGAGAGAACTACGAAAAAGTCAGAGCT	2072
QY	1999	CAGACACAGACTCAAGAAATACGTCTTGATATTACCCACGAAAGTGTGTTAACCGAA	2058
Db	2073	GAAGAATGTACGGCGAGTACGTGTGTGACGTAAATAAAAATGCTCGTCMAAGAGAG	2132
QY	2059	GACCGAGGTCCCTGTGTGCTTAACCGGTGATGTGTGTAATCCACCATTTTCAAGATTGCG	2118
Db	2133	GAAGCTTCGGGTTTGTGTGTGTGTGTGGAGAGTTAACCAACCCCGCTTCATGAATTGCGC	2192
QY	2119	TACGAGACTCTCAAGACACGACGACGAGCACTTCAAAAGTCCCAACATCGAGTCTAT	2178
Db	2193	TACGAAAGGGCTGAAGATCAGGCGCTCGGACCATATTAAGACTACAGTATGAGAGTCTTT	2252
QY	2179	GGAGTCCAGGTTCAAGTAAATCTGGAATCATCAAAAGCCTGTGACTAAGAAAGATCTG	2238
Db	2253	GGGGTTCGGGATACAGCAAGTCTGTCTATTAATTAAGAGCTCTCGAACCAACACGATCTG	2312
QY	2239	GTTGTGAGTGGGAAGAAAGAAAACTGGCGCAAAATCATCGGCACTGTAAAGAGATGAGA	2298
Db	2313	GTCAACACCGGCAGAAAGAGAACTGCCAGAAATAGTTAACGAGTGAAGAACACCGC	2372
QY	2299	CGTATGATGTGTGCTAGGACTGTGCATTCAGTGTCTTAAATGGGGTTAAAGCAACCC	2358
Db	2373	GGGAAGGGGACAAGATAGGAAAAACAGTGACTCCATCTCGCTAAACGGGTGTGCTGTGCC	2432
QY	2359	GTAAACACTCTGTACATTTGATGAGGCAATTCCTGCCATGACGAGGACGCTGTGGCACTG	2418
Db	2433	GTGACATCTCATATATGTGACGAGGCTTTCGCTTCCATCCGGTACTCTGTGGGCCCTTA	2492
QY	2419	ATTGCGATTCGCAAACT---AAGAAAGTGGTATTGTGGCGGGACCCAAAACAATGGGCG	2475
Db	2493	ATTGCTCTTGTTAACTTCGGAGCAAGGTGTGTTATGCGAGACCCCAAGCAATGGCGA	2552
QY	2476	TTCTTTAACATGATGTGCTGAAGATGATCATTTTAACATGACATATGACATGAGTGTAC	2535
Db	2553	TTCTTCAATATGATGACAGCTTAAGTGAATTTCAACACAACTCTGCACTGAAGTATGT	2612
QY	2536	CATAAAGCATCTCTAGAGGTGACACAGACTGTAAACCGGCATCGTCTCACGCTCTTC	2595
Db	2613	CATAAAGTATATCCAGACGTTGACCGCTGCATCCAGGCATCGTCTTAAGCTTGCAC	2672
QY	2536	TACGACAAAGCAATGAAGACGGTTAAACCCATGTGCTATAAATCATCATATGATATCCACA	2655
Db	2673	TACGAGGCAAGATGCGCACGACCAACCCGTGCACAAACCCCATATATATGACACACA	2732
QY	2656	GGGACCAAAAGCCGCAAAAGATGATCTGATTTTAACCTGTTCACAGAGTGGGTGAAA	2715
Db	2733	GGACAGACCAAGCCCAAGCCAGAGACATGTGTAAATATGCTTCGAGGCTGGGCAAG	2792
QY	2716	CAGCTACAGATTGACTACAAAATACAGAAATCAATGACTGCGGCTGCACAGAGACTT	2775
Db	2793	CAGCTGACGTTGACCTACCGTGGACACGAAATCAATGACAGACAGCATCTCAAGGCGCTC	2852
QY	2776	ACGGGAAAAGCGCTTATGTGTGTCAAGSTCAAAATCAACGAAATCCAATCTACTCGCAG	2853
Db	2853	ACCGCAAAAGGGGTATACCGCGTAAAGGCAAGAGGTGATAAAAATCCCTGTATGCCCT	2912
QY	2836	ACTTCTGAGCACTGAAGCTGTACTTACACAGCAACAAAAAACGCAATTCCTGAAGACG	2895
Db	2913	GCCTCGAGACACGTGAATGTATCTGCTACCGGCACCTGAGGATAGCGTGTGTGAAAAACG	2972
QY	2896	CTACTGCTGATCCCTGATTAAGACATTACAGCTAAATATCCCGGGAGTTTCACGCT	2955

Db 2973 CTGCGCCGCGATCCCTGGATTAAAGTCTCTATCAAAACATTCACAGGGGTAATTTACGGCC 3032
Qy 2956 TCATTGAGCACTGGACGCGGAAACGACGCCATTTATGGCAGCGCTTTGATTAAGCCG 3015
Db 3033 ACATTGGAAAGATGGCAAGAAACACGCAAAATTAATGAAGTGATTAAGGACCGGCT 3092
Qy 3016 CAGACAGCTGATGTGTTCCAGAAATTAAGGTGAACGCTGCTGGGGGAAAGGCTTTAGAGCA 3075
Db 3093 GCGCCTGTGACGCGTTTCAGAACAAACGAACTGTGTTGGGGAAAGCGCTGTGCT 3152
Qy 3076 GTCTTGGCAGCGCCAACTATGTGCTGACGACAGCAGTGGAGAC---GTTGACCCCA 3132
Db 3153 GTCTGGCACTGCGCGAATCAGATTGACAGCAGAGGTGGAGACCATTAATTAACGCA 3212
Qy 3133 TTCAAGCATGACAGAGCTACTACCTGAAATGGCACTGAACCTTTTGGACCAAGGCTC 3192
Db 3213 TTTAAGGAGCACAAGCTTACTCTCAGTGGTGGCTTGAATGAATTTGACCAAGTAC 3272
Qy 3193 TTTGAGTAGACCTGGACAGTGGGTTATTTCCGCTCCTACCGTCCGCACTTACTTACAG 3252
Db 3273 TATGAGTAGACTGGACAGTGGCCTGTTTCTGCCCCGAAAGTGTCCCTTATTTACAG 3332
Qy 3253 GATCAGCACTGGGATTAATCCGCCAGGGAAGAAATGTATGGCTTAATAGAGGTAGCA 3312
Db 3333 AACACACACTGGGATTAACAGACCTGTGGAGATGTATGATTCATGCCGCAACAGCT 3392
Qy 3313 AAGAGTTGTACGCGCATATCCGTGATCACAAAAGCGTTGACACAGGAGGTAGCT 3372
Db 3393 GCCAGGCTGGAAAGTAGACATACCTTCTGAAAGGGGAGTGGCATCCGGGCAACAGGCA 3452
Qy 3373 GATTAAGGAATTAATACATCAAGGACATCTCCAAATTAATGTGGTCCATTAAT 3432
Db 3453 GTTATCCGAGAAAGAAATCCAAACCGCTTTCTGTGCTGACAAATGTAATTCATATCAC 3512
Qy 3433 CGCCGTTGCCCACTCGTTGATGTTGACCAAAAGGACAGGGTACAACTGATCACAGC 3492
Db 3513 CGCAGGCTGCGCAGCGCCCTGTGGCTGATGTAACAAGACGTTAAAGGAGTAGGTTAG 3572
Qy 3493 GGATTCCTATTAAGATGAAGGCAATCTGTGTGATGCGGATCTTATCAGATT 3552
Db 3573 TGGCTGTGCAATTAAGTAAGAGGGGTACCACTCTGTGTGATGAGTACCACTGGCT 3632
Qy 3553 CCAGGAAGAAAGTAAGATCCATGGGTCCATTGCCCACTAATAC-----CATCAG 3603
Db 3633 TTGCTCGACCGGACGTCATTTGTTGTACCGCTGAATGTCACGCGCCGATAGGTGC 3692
Qy 3604 TGTGATCTGATTTGGGAATACCTAGCCATGTGGTAAATATGACATTAATCTTTGTCAAT 3663
Db 3693 TACGACCTAAGTTAGACTCCGCGTGAACGCGGCAAGTTGCACTTGTGTTGTGAAAC 3752
Qy 3664 GTTAGAGCCCGTACAGAACCATATCAACCAAGTGGAGATCAGCTATCCACCAC 3723
Db 3753 ATTACACGGAATTCAGATTCACCACTACAGCAGTGTGTGACCAAGCCATGAAAGTGT 3812
Qy 3724 AGCANTGTAAGTGAAGTGTGTCACCACTGAAACATGCGGGAACATGTGGCTATA 3783
Db 3813 CAGATGCTGGGGAGATGCGCTACGACTGTAAACCCGGGGGATCT---TGATGGA 3869
Qy 3784 GGGATGAGGCTTGTGATCGCGCAACCGAATATCATCTGCGGTGGACGCTCATTT 3843
Db 3870 GCTTACGATACGCGGATTAATTAAGCGAAGCCGTTGTTCTCTTAAAGCAAAAGTTC 3929
Qy 3844 AGGTTTACCGGTGTGTGACGCTTAAGAACCTGCGGAAATACTGAGGTTCTTCTG 3903
Db 3930 TCGTCTGCAAGAGTGTGCGCCGATGTGTGACCAAGATTAACAGAGTCTTCTGCTG 3989
Qy 3904 TTTCTTGGCAAGGACAAGGCAACACACATGACCGAGCAGAGCTGGGTAGTGCTT 3963
Db 3990 TTTCTCAACTTTGACAAAGGAAAGAGACCTCTTACGCTACCAAGATTAATCAAGCTG 4049
Qy 3964 GACAACATCTATCAAGGCTCAACAGGTACGAGGAGGAGAGCTCCAGCTTACAGAGTG 4023

Db 4050 AGTCCGCTATGCGCGAGGAAGCCATGACACGCGCGGTGTGACCATCTACAGAGTT 4109
Qy 4024 ATCAGAGTGAATTTACAGAGCGCTGACCAAGCTATGCTTAATGCTGAATAGGAAA 4083
Db 4110 AAGAGACAGACATAGCCACGTGACAGAACCGCTGTGTATACGAGCTAACGCCCT 4169
Qy 4084 GGTCACACAGATTCGAGAGTGTGGGTGACCTGTATCCGAAATATGCGCGGTCTTTTAT 4143
Db 4170 GGAATGTAGGGAGATGGCGTATGAGGGCCGTGGCGAAAGAAATGGCCGTCAAGCTTAAG 4229
Qy 4144 AGACAGCAATAGTGTGCGGACGCTGACCTGTGAGACGACCAAGC---CTCATATA 4200
Db 4230 GAGACACCAACACAGTGGGACACATTAATAACAGTCAATGTGCGGCTGTATCCCGTATC 4289
Qy 4201 CAGCTGTAGGACCCATTTTCTTAAGATCCGGAACCGGAGGCGCACTTAAGCTGCA 4260
Db 4290 CACGCTAGCGCTTAATTTCTGTGCAACATGAAACGGAAGGGAGCCGGAATTGGCC 4349
Qy 4261 GCTGCTACATAGACATAGCGTCCATGTCACACGCTGAGCGGATTAACAAAATATCAGTA 4320
Db 4350 GCTGTCTACCGGAGAGTGCCTCCGGAAGTAAACAGACTGTACAGACAGGTAGCCATC 4409
Qy 4321 CCGCTACTGTCAACCGGCATCTATTCTGTGGCAAGATGAGTATGCAATTCGAT 4380
Db 4410 CCGCTGCTGTCCACAGAGAGTTCAGCGCGGAAAGATAGGCTGACAGCAATCCCTCAAC 4469
Qy 4381 CACCTGTTCATGCTTTTCGACATGAGGATCCGATGTCACATATATGCTTGGATTA 4440
Db 4470 CATTTATTCACAGAAATGAGACGACGACGCTGACATGACATCTACGACAGACAA 4529
Qy 4441 CAATGGAGACACAGATTAATGAGGCAATTCACCGCAAAAGAGCTGAAATTCGAT 4500
Db 4530 AGTTGGAGAAAGAAATTCAGAGAGCCATTGACATGAGAGCGGTGTGAGTTGCTCAT 4589
Qy 4501 GATGACAAAGCCAGTGAATTTGACTTGTGATGAGGTCCAAACCAACAGCTCTTGGCAGC 4560
Db 4590 GATGACCTGAGGCTGACCAACAGACTTGTGATGAGTGCACCCGACACAGCGCTGTGGGT 4649
Qy 4561 AGACCAAGTAACTCCGTCATGAGGGGCAAGTTGATTCATACCTGGAAGGTACAGATT 4620
Db 4650 CGTAAAGGCTACAGTACCACTGACGAGGTCCGTGATCTGTAATTTGAAGGTACGAATTC 4709
Qy 4621 CATCAGACCGCAGAGCAATTTGCGGAATCCATGCAATGTGGCCCAACAAATCTGAGGT 4680
Db 4710 AACGAGGCTGTATGATTAAGGACAGATTAAGTGTGAGTGTGAGCCCAACAGCAAGGCA 4769
Qy 4681 AATGACGATTTGCTTGTATCATCTGTGGGAGATATGTCCAGATCCGCTCCAATTC 4740
Db 4770 AACGACAGATATGCTATATGACGCTGTGGGCAAAACAAATGGAACAATCAGATCCAAATGT 4829
Qy 4741 CCAATAGAGATGAGAGGGGTGCTCCACCTGACACACTTCATGCTGTGTATTAAC 4800
Db 4830 CCGGTGAACGATTCGATTATCAACACTTCACAGACAGTGCCTGTGCTGTGCGCTAC 4889
Qy 4801 GCTATGACGCGTGAAGCCGATATACAGTTGCGCTGTGCGGAAGAAAGAGTGTGCGGTA 4860
Db 4890 GCAATGACAGAGAACGATTCGCGGCTTAAGTCAACACAGTAAAGATGATGTGTT 4949
Qy 4861 TGCTCATCATTTCTGTTGCCGAGTACAGATCAACAGCGGTGCGAAGCTACAGTCAAC 4920
Db 4950 TGCTCATCTTTTCCCTCCGAAATACATGTAATGAGGGGTGCGAAGGTAAATGCGAG 5009
Qy 4921 AACACAGTCTGTTTACAGGCGTGTACACCGGCTTATACCCCGAAGAGTACGCGGA 4980
Db 5010 AAGGTTCTCTGTTGACCCGACGATCTTCACTGTGTAGTGTGCGGAAAGTATGCGCA 5069

RESULT 7
US-09-901-106-1
; Sequence 1, Application US/09901106
; Patent No. US20020151067A1
; GENERAL INFORMATION:
; APPLICANT: Garoff, Henrik

```

1  Ilijestrom, Peter
2  TITLE OF INVENTION: DNA Expression Systems Based on
3  Alphaviruses
4  NUMBER OF SEQUENCES: 27
5  CORRESPONDENCE ADDRESS:
6  ADDRESSEE: Birch, Stewart, Kolaesch & Birch
7  STREET: P.O. Box 747
8  CITY: Falls Church
9  STATE: Virginia
10 COUNTRY: USA
11 ZIP: 22040-0747
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE: Floppy disk
14 OPERATING SYSTEM: IBM PC compatible
15 SOFTWARE: Patentin Releasee #1.0, Version #1.25
16 CURRENT APPLICATION DATA:
17 APPLICATION NUMBER: US/09/901,106
18 FILING DATE: 10-Jul-2001
19 CLASSIFICATION: <Unknown>
20 PRIOR APPLICATION DATA:
21 APPLICATION NUMBER: US/07/920,281C
22 FILING DATE: 13-AUG-1992
23 ATTORNEY/AGENT INFORMATION:
24 NAME: Murphy Jr., Gerald M.
25 REGISTRATION NUMBER: 28,977
26 REFERENCE/DOCKET NUMBER: 828-103P
27 TELECOMMUNICATION INFORMATION:
28 TELEPHONE: 703-241-1300
29 TELEFAX: 703-241-2848
30 TELEX: 248345
31 INFORMATION FOR SEQ ID NO: 1:
32 SEQUENCE CHARACTERISTICS:
33 LENGTH: 11517 base pairs
34 TYPE: nucleic acid
35 STRANDEDNESS: single
36 TOPOLOGY: linear
37 MOLECULE TYPE: RNA (genomic)
38 HYPOTHEITICAL: NO
39 ANTI-SENSE: NO
40 ORIGINAL SOURCE:
41 ORGANISM: Semliki Forest Virus
42 FEATURE:
43 NAME/KEY: -
44 LOCATION: 1..11517
45 OTHER INFORMATION: /label= genome
46 /note= "Semliki Forest Virus complete nucleotide
47 sequence, presented as a cloned DNA sequence; see
48 Figure 5."
49 FEATURE:
50 NAME/KEY: CDS
51 LOCATION: 87..7379
52 OTHER INFORMATION: /product= "SFV polypeptide"
53 FEATURE:
54 NAME/KEY: CDS
55 LOCATION: 7421..11179
56 OTHER INFORMATION: /product= "SFV polypeptide"
57 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
58 US-09-901-106-1
59
60 Query Match 31.3%; Score 1564; DB 10; Length 11517;
61 Best Local Similarity 58.3%; Pred. No. 0;
62 Matches 2904; Conservative 0; Mismatches 2040; Indels 36; Gaps 8
63
64 Oy 31 AGAATTCAGTTGACTAGATGCTGACAGCCCGTATGTCAGTGGTACAGCGGACGTTT 90
65 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
66 Db AAAGTGCATGTTGATTAATGAGGCTACAGCCCATTCATCAAGTCTTTGCAAGGCAATTT 155
67 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
68 Oy 91 CCACAATTGAGATGAAGCAAGGACAGGCTACTGACAAATGACCATGACCATGCCAGAGCG 150
69 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
70 Db 156 CCGTGTTCGAGGTGAGTGCATTCGACAGGTCAACCAAAATGACATGCAAAATCCAGAGCA 215
71 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
72 Oy 151 TTTTCGATGTGGCAACAAGCTCATTTAGAGACGGAATGACCCGGAGCAACTTATCTTGG 210
73 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Db	216	TTTTGGCACTGGCTACCAAAATTGATCGACGGAGACTGACAAAGACACTCATCTTG	275
Qy	211	GACATTGGAAAGTCGCGCCCGTCAGACATGCACATTCCAATTCACCGGTATCATTTGATCTGC	270
Db	276	GATATCGGCAAGTGGCGCTTCCAGGAGAAATGATGTCTACGCAAAATPACACTGCGGATGCG	335
Qy	271	CCTATGATAGCGCTGAAGACCCGGACAGACTACACCGTATGCGAAAGACT-----T	324
Db	336	CCTATGCCACAGCGCGAAGAACCCCGAAAGGCTCGATAGCTACGGCAAGAACTGGCAGCG	395
Qy	325	AAGAAATGACATTACCGACAAAGAACATAGCCTTAAGGGCGGACCTGCTGGAAGTC	384
Db	396	GCCTCGGAAAGGTCTGGATAGAGATCGAGAAATATCACGACCTGACAGCCGTG	455
Qy	385	ATGTCAACACCGACGACGAGACCTTCATCTCTGTATATGACACAGACGCCACGTGTAG	444
Db	456	ATGGCTACGCCAGACGCTGAATCTCTCACTTTGCTCGCATACAGACGTCAAGTGTGCT	515
Qy	445	TACTTTGGAAGTGTAGCAGTATACCAAGATGTAGCGCATGCATGACCGACATCAATC	504
Db	516	ACGGCAGCGGAAGTGGCCGTATACCGAGACGTATGCTGTACATGACCAACATCGCTG	575
Qy	505	TACCAACGAGGGCTTAAAGAGTTAAGGCAATTTTACGTAGGGCTTTGACACGACCCCT	564
Db	576	TACCATCAGGGAGATGAAGGTGTCAAGACGGCGTATTGGATTTGGTTGACACACCCG	635
Qy	565	TTTATGTACAAAACATGCGAGGTTCTTCACTTACTTTCACACGAACTGGGCTGACGAG	624
Db	636	TTTATGTTTGAAGGGCTAGAGGGCGCGTATCCAACTTAGCCACAACTGGGCGACGAG	695
Qy	625	AGAGTATTGGAAGCACGTAACTATGGCTCGGTAACTAGATCTTCAGAGAGCAGCTT	684
Db	696	CAGGTGTATACAGGCCAGGAACATAGGACTGTGTGACGATCCTTACTGAGGGAAGACTC	755
Qy	685	GGAACACTCTCAATCTTAGAGAAAGAGGCTCCAACTTACTTAATTAATCATTTCTG	744
Db	756	GGCAAACTGTCATTCTCCGAGAAAGAAATTGAAACCTTGGCACAGTCAATGTTCTCG	815
Qy	745	GTTGGTTCAACAATCTTACACAGAAAGATGATATCACTGTACGTAGCTGCATTTCCAAAC	804
Db	816	GTAGAGTCTACATGTGTACACTGAGAGCAAGAAAGCTATCTAGGAGCTGGCATTTACCTCC	875
Qy	805	GTTGTTCCACTTGAAAGAAAGTCTTAATCTTCAAGGTAGATGTGGAGCAATGTGCAGCTGT	864
Db	876	GTAATTCACCTGAAAGGTAAACAATCTTTACCTGTAGGTGGATTAACATCGTATCATGT	935
Qy	865	GAAGGTTACGTATCAAAAAGATACGATCAGCCCAAGACTATACGGTTAAAGTTAGAAAC	924
Db	936	GAAGGTAACGTAGTTAAAGAAATCACTATGTGCCCCCGGCTGTAGGTAAACGGTATAGG	995
Qy	925	TTGGCGTCCAAATGATCGCGAGGGTTCTTGAAGTTCAAAAGTCACAGATACGCGTCGC	984
Db	996	TACGCCGTGACGTATCACGCGGAGGGATTTCTTAGTGTCAAGACACACAGACTGTCAAA	1055
Qy	985	GGCGAGAGGGTTTCTTTGCTGTGTGTACGTATGTACCAAGCCACACTTTCGATCAGATG	1044
Db	1056	GGAGAAAGATCTCATTTCCCTGTATGACCTACGTCCTTCAACATCTGTGTCAAAATG	1111
Qy	1045	ACAGGGATTTCTGGCACTGACGTTTAGTGTGATGACGCACAAAAATCTATTGTTGGCTC	1104
Db	1116	ACTGGCATCTAGAGGACCGACGTACACCGGAGGACGCAAGMACTTGTATGTAGGATTTG	1177
Qy	1105	AACCAAGGATTTGCTCAATGTGTAGAGCGCAAGAAATATCTAACACATGAGAAACTAT	1166
Db	1176	AATCAGAGGATAGTTGTGACGGAAGAACACAGCGAAACCTTACACGATGAAGAACTAT	1233
Qy	1165	CTATTACAGTGTGCGCCAGCGCTTTTCAAGTGGGCGCGTGAACATCTGTCGCACTTG	1222
Db	1236	CTGCTCCGATTTGTGGCGCTGCAATTTAGCAAGTGGGAGGAAATACAAAGCAGACCTT	1299
Qy	1225	GACGACGAAGAACTTAGGGGTGCGGAGCGCACTTTACTATAGGGCTGCTGGGCT	1288

Dh 1296 GATGATGAAAAAAGCTTGGGTGTCCGAGAGAGTCACTTACTGTGCTGTGGGCA 1355
Qy 1285 TTCAAGACCCGAAATACATCCATCTACAAAGAAAGCTGTAGCGCAAACTTAAGAA 1344
Db 1356 TTTTAAACGAGAAATGACACACATGTACAAAGAAACCAACACCAATATGTAGAG 1415
Qy 1345 GTACCTGCGGCTTTGATCTATTGTGATTTCCAGGCTTACCGACCAAGGCTGATATG 1404
Db 1416 GTGCTTCAAGATTAACTGCTGTGATCCCGAGGCTATGTGTACAGGCTTGGCAATC 1475
Qy 1405 GCGTTCGCGCGTGAAGCTGCTGCTGTAACCACTGTCAAAACCGGCTATTT 1464
Db 1476 CCACTCAGATCAACCATTAAGATGCTTTTGGCAAGAAAGCAAGC---GAGATTAATA 1532
Qy 1465 ACATGCGCCGATGTGAGCATCTGCTGAGCTTAACAGAAAGCTGAAGAGTGGCTGCA 1524
Db 1533 CTGTCTTCGACGGGTGTCGTCAGCGAGGATGTGAACCAAGAGGAAAGAGATGGAG 1592
Qy 1525 GCGGAGAGATCAGAGAAAGCCCTGCAACCTTGTCTCCCTGAATTAAGAAAGAGACGTA 1584
Db 1593 GCCGAGCTGACTAGAGAAAGCTTACACCCCTCGTCCCATCGCGCGCGAGACGGGA 1652
Qy 1585 GAGGCAAGATGACCTCTATTATGCAAGAG-----GCAAGAGAGATGAGCTGAGACA 1638
Db 1653 GTCTTCACAGCTGACGTTGAAGAACTAGAGTATCACGACAGGTGCGAGGGGTGTGAAACA 1712
Qy 1639 CCAGAGAGACATCAGGGGTGACAAAGTTACCCAGCGCAAGAAAGATGGGTCTTAGCCT 1698
Db 1713 CCTGCAAGCGGCTGTAAGATACCGCACAGCGCAAGCACTACTAGAAATTAAGCTA 1772
Qy 1699 ATACTTTCACCCCAAGCGGCTATTGAATAGTAAAGAACTGCGGTATTCACCCATTGGCG 1758
Db 1773 GTTGTGTCCTCCGAGACCGTGTCTCAAGAGCTCAAGTTGGCCCCCGTGCACCTCTAGCA 1832
Qy 1759 GAACAAGTACTGTATGACTCACAAAGTAGGGCAGGAGATCAAAATGTGAGCCATAC 1818
Db 1833 GAGAGGTGAATAATATACATATACGAGAGCGCGGCGGTACAGATGAGCGATAT 1892
Qy 1819 CAGCGTAAAGTCACTTGTACCGAAAGGAGCGCGGCTCCCTGTTCAAGACTTCCAGCATATG 1878
Db 1893 GACGGCAGGGGTCTACTACCACTGTGATCGGCAATTCGGTCTCTGAGTTTCAAGCTTTG 1952
Qy 1879 AGTGAAGCGCTGATCGATGTTTTCAACGAGAGGAGTTCTGAACAGATACCTGACACAC 1938
Db 1953 AGCGAGAGCGGCACTATGTGTACACGAAAGGAGGTTCTGTCACAGAACTATATCCAT 2012
Qy 1939 ATTCGAATCAACGAGAGAGCGCTAAACACTGACGAAGATATCTAATAGACTGAAGACT 1998
Db 2013 ATTGCGGTTCAAGGACCGTGTGTAACACCGACGAGAGGAACTACGAAAGTGCAGAGCT 2072
Qy 1999 CAGGACACAGCTGAAATAGTCTTGATATGACCGACGAAAGTGTGTTAACCGGAA 2058
Db 2073 GAAAGAACTGACGCGAGTACGTGTTCAGATGAATAAATACTGTGTAAGAGAGAG 2132
Qy 2059 GACGAGAGTCCCTTGTGCTTAACCGGTGATGTGATGATCCACATTTCAAGATTTGCG 2118
Db 2133 GAAACGTGCGGTTGTGTGTGTGTGAGAGAGTAAACCAACCCCGTTCCAGATTTGCC 2192
Qy 2119 TACGAGAGTCTCAAGACACGACGAGCACTTCAAAAGTCCCAACCATGAGTCTAT 2178
Db 2193 TACCAAGGCTGAGATCAGGCGCTCGGACCATATAGACTACAGTATGAGAGTCTTT 2252
Qy 2179 GGAATGCGAGTTGATAGTAAATCTGGAATCATCAAAAAGCGTGTGACTAATAAAGATCTG 2228
Db 2253 GGGGTTCCGGGATAGGCAAGTCTGTATTAATTAAGAGCTCTGACCAAAACAGATCTG 2312
Qy 2239 GTTGTGATGCGAAGAAAGAACTGCGCAAGAAATCATCAGGAGTGAAGAGATGAGA 2298
Db 2313 GTACACAGCGGCAAGAGAGAACTGCCAGAAATAGTAAACGACGTGAGAGAACACCGC 2372
Qy 2299 CGATAGATGTTGTGCTGAGAGCTGTGATTCAGTGTCTTAAATGGGGTTAAGCACCC 2358
Db 2373 GGGAAAGGGAACAAGTAGGAAACAGTGAATCTCATCTGTAAACGGGTTGCTGTGCC 2432

Qy 2359 GTTAACATCTGTACATTTGATGAGGATTTTCCCGCATGACAGGAGCGTGTGGCACTG 2418
Db 2433 GTGACATCTCTAATGTGTGAGAGAGCTTTGCTTGGCATTCGGTACTCTGTGAGCTTA 2492
Qy 2419 ATTGCACTGTCAAACT---AAGAAAGTGTATTTGTGCGGGGAGCCCAAAACAATGGCGC 2475
Db 2493 ATTGCTCTTGTAACTCGAGCAAAAGTGTGTATGCGGAGAGCCCAAGCAATGGCGA 2552
Qy 2476 TTCTTTAATCATGTATGCTCTGAAAGTACATTTTAACATGACATATGACACTGAAGTAC 2535
Db 2553 TTCTTCATATGTATGACGCTTAAGTGAACCTTAACCAACAATCTGCACTGAAGTATGT 2612
Qy 2536 CATTAAGACATCTCTAGAGGTGACACAGACTGTAAACCGCATGTCTGCAAGCTCTTC 2595
Db 2613 CATTAAGTATATTCAGACGTTGACGCGCTCACTCAAGCATGTGTCTACGTTGCAC 2672
Qy 2596 TACGACAAAGCGAATGAAGAGGTTAAACCATGTGCTGATTAATCATATAGATACCA 2655
Db 2673 TAGGAGGCAAGATGGCCAGACCAACCGTGCACAAACCAATATATAGACACCA 2732
Qy 2656 GGAACCAAAAGCCGCAAAAGATGATCTATTTAACTGTTTCAAGAGATGGGTAAA 2715
Db 2733 GGAACGACAAAGCCCAAGCCAGAGACATGTGTAAATGCTTCCAGGCTGGCAAG 2792
Qy 2716 CAGCTACAGATTTGACTCAAAATATCAAGAAATCATGACTCGGCTGATGCAAGACTT 2775
Db 2793 CAGCTGAGTTGAGACTACCGTGACAGAGATGATGACAGACAGATCTCAGGCGCTC 2852
Qy 2776 ACAGGAAAGCGGTTATGTGCTGACAGTACAAAGTCAAGAGATCCACTCTACGCA 2835
Db 2853 ACCCGCAAGAGGATATACCGCTGAAGCAGAGATGAATGAATATCCCTGTATGCCCT 2912
Qy 2836 ACTTCTGACAGCTGAACGTTACTTACTACGACAGCAAGAAAGGATTTCTGGAAGAC 2895
Db 2913 GCGTCGAGACGATGATGTATGCTGACCGGACGCACTGAGATAGCGTGTGGAAGAAC 2972
Qy 2896 CTAGCTGTATCCCTGATTAAGAAAGACCTTACAGCTAAATATCCCGGGATTTCAAGGCT 2955
Db 2973 CTGGCCGCGATCTCTGATTAAGGTCTTATCAAACTTCCACAGGTAATCTTACGCGC 3032
Qy 2956 TCAATGACACTGTGACGCGCAACAGACGCACTTAATGCGCGCTTCTTGAATAGCCG 3015
Db 3033 ACATGGAAGATGCGCAAGAAAGAACAGCAAAATATGAAGTATGAAGACCGGCT 3092
Qy 3016 CAGACAGCTATGTGTTCCGAATTAAGTGAAGTCTGCTGGCGGAAGGCTTATGAGCCA 3075
Db 3093 GCGGCTGTGACGCGTTCCGAACCAAGCAAGCTGTGTGGCGCAAAAGCTGTGTCCT 3152
Qy 3076 GTCTTGCGACGCGCAACATTTGTGCTGACGAGACAGCAGTGGGAGAC---GTTGCAACCA 3132
Db 3153 GTCTGAGACCTGCCGAATCAGATTTACAGAGAGTGTGAGACATATATTAACGCA 3212
Qy 3133 TTCAAGCATGACAGAGGTACTGACCTGAATGAGCACTGAATCTTTTGCACCAAGTTTC 3192
Db 3213 TTTAAGGAGACAGAGCTTACTCCAGTGTGCGCTTGAATGAATTTTSCACCAAGTAC 3272
Qy 3193 TTTGAGTAACTCTGACATGTGGTTATTTTCCGCTCTTACCGTGCACCTTAATAGG 3252
Db 3273 TATGAGATTAACTGTGACAGTGTGCTGTTTTCGCCCCGAAGTGTCTGTATTAAGAG 3332
Qy 3253 GATCAGACTGGGATTACTGCGCAGGGAAGAAATGATGAGGCTTAAATAGAGAGTGA 3312
Db 3333 AACACCACTGGATTAACACACTGTGTGAGAGATGATGATTTCAATGCGCAACAGCT 3392
Qy 3313 AAGAGTTGTACGCGCATATCTGTGATCAAAAAGCGGTGAACAGGCAAGGATGACT 3372
Db 3393 GCAAGCTGGAAGCTAGACATCTTCTGAAGGGGAGTGGCATACGGGCAAGCAGCA 3452
Qy 3373 GATTAAGAAATATATCAACAGACTACTCTTCCAACTTAATTAATGTGTGTTCAATTAAT 3432
Db 3453 GTATATCCAGAAAGAAATATCAACCGCTTGTGTGCTGAGCAATGATTAATCTTATCAC 3512

Db 2882 CCTATGCGCAGCGAGAAAGACCCGAAAGGCTCGATAGCTACGCAAGAAATCGGACGG 2941
Qy 325 AAGAAAAGTGCATTACCGACAAAGAACTATACCTCTTAAGCGCGAGACCTCTGGAATC 384
Db 2942 GCCTCCGGGAAAGGCTGAGATAGAGAGATCCAGAGAAATACCGACCTTCAGACCGCTC 3001
Qy 385 ATGTCAACACCAAGACGAGACCTCCATCTCTGTATGACACAGACGCCACGGTATAG 444
Db 3002 ATGGCTACGCAACGCTGAATCTCTTACCTTTGCTGCACTACAGACGTACGTGTCTG 3061
Qy 445 TACTTGGAAAGTATAGACATATACCAAGATGTATACGAGTCAATGCAACGACATCAATC 504
Db 3062 ACGGACCGGAAGTGGCGGTATACAGGACGATATGTCTATACATGACCAACATCGCTG 3121
Qy 505 TACCAACAGCGGCTTAAAGAGTTAGCAATTTACTGTGATAGGCTTTGACACGACCTT 564
Db 3122 TACCATCAGCGGATGAAAGGTGTCAAGACGCGATTTGGATTGGATTGACACACCCCG 3181
Qy 565 TTTATGTACAAACATGCGAGGTTCTTACCTACTTACACAGAACTGGGCTGACGAG 624
Db 3182 TTTATGTTGACGGCTAGCAGCGCGGTATCAACTACGCAAACTGGGCGAGCGAG 3241
Qy 625 AGAGTATGGAAGACGTAACATTGGCTCGTAACTCAGATCTTCAGAGAGAGAGGCTT 684
Db 3242 CAGGTGTACAGGCGAGGAACATAGAGACTGTGTGAGCATCTTGCATGAGGAAGATC 3301
Qy 685 GGAAGAACTCTCAATCTTATAGAGAAAGGCTCCAACTTAAATAGATATATTTCTG 744
Db 3302 GCGAAACGTCTCATCTCCGAGAAAGCAATTTGAACCTTCGACACAGTATGTCTCG 3361
Qy 745 GTTGTATCAACATCTACAGAAAGATAGATCACTGTACAGTAGTGATCTTCCAAAC 804
Db 3362 GTAGATTTACATTTGTACACTGAGCAGAAAGCTACTGAGAGCTGCACTTACCTCC 3421
Qy 805 GTGTTCATTGAAAGAAAGTCTTAATTCAAGGTAGATGTGGAGCCATTGTACGCTGT 864
Db 3422 GTATTCACCTGAAAGGTAAACAATCTTTACCTGTAGTGAGATACCATGATCATGT 3481
Qy 865 GAAAGGTCTCTATCAAAAAAGTAAAGTACATGACCCAGACATATACGTTAAAGTTGAGAC 924
Db 3482 GAAAGGTACGTATGTAAAGAAATCACTATGTGCCCGGCTGTACGTTAAACGGTAAAG 3541
Qy 925 TTGGGTCCACAATGCAATGCGAGAGGTTCTTGAAGTTCAAAGTACAGATACGTCGCG 984
Db 3542 TACGCGTACGATACGCGAGAGGATTCCTAGTGTCAAGACCAACGACATCTGTCAA 3601
Qy 985 GCGGAGAGGTTCTTTTGTGTGTGTATGATATGACGACGACATTTGGCATGATG 1044
Db 3602 GGAGAAAGATCTATTCCTGTATGACCTACGTCCTCAACATCTGTATCAATATG 3661
Qy 1045 ACAGGAGATTCTGCAATGAGTATGTGATGACGACAAAAATATTTGGTTGGGCTC 1104
Db 3662 ACTGGCATATACGACCGACGATCACACGAGAGGACGACAGAAAGTTCTTGTGGGATG 3721
Qy 1105 AACCAAGAGTTGTGTCAATGTGTAGAGCGAAAGAAATCTAACACAAATAGAGAACTAT 1164
Db 3722 AATCAGAGATATGTTGTAAAGAAACACGCAAAACCTAACACATAGAAATAT 3781
Qy 1165 CTATTTACAGTGTGCGCCAGCGTTTTCAGGTGGGCGGTGAACATGTGCGCATTTG 1224
Db 3782 CTGCTTCGATTTGTGGCGGTGCGCATTTAGCAAGTGGGAGAGGAATACAAAGCACTT 3841
Qy 1225 GACGACGAAAGAACTAGGGGTGCGGAGCGCACTTTACTATGCGCTGTCTGGCT 1284
Db 3842 GATGTATGAAAAACCTCTGGGTGTCCGAGAGAGGTCACTTCTGCTGTGGGCA 3901
Qy 1285 TTCAAGACCCAGAAATCAATCATCTTACAAAGAAAGCGGTAGCGCAAAATTAAGAA 1344
Db 3902 TTTAAAGAGAGAGATCAACCATGTACAAAGAAACAGACCCCAACATATAGTAG 3961
Qy 1345 GTACCTGCGCTTTGACTCATTTGTGATTCACCGCTTACAGGCAAGCGGCTGATATG 1404

Db 3962 GTGCTTCAAGATTTAATCGTTTCATCCGAGCCTATGCTTACAGGCTTCGAATC 4021
Qy 1405 GGTTCGCGGTATGGCTCAGCTGCTGTGAACCAATGTCAACCCGACCGGCTATT 1464
Db 4022 CAGTCAATACGATTAAGATGCTTTTGGCCAAAGAACCAAGC--GAGATTATA 4078
Qy 1465 ACAATGCGCATGTGAGCATCTGCGTGTCTTACAGCAAGACCTGAAGAGTGGCTGA 1524
Db 4079 CTGTCTCGACCGGTGTGAGCAGGAGATGTGAACAAAGAGAAAGAGAGTGGAG 4138
Qy 1525 GCGGAAGATACAGAAAGCTTGCACCTTGTCTCTTGAATTAAGAAAAAGACCTGA 1584
Db 4139 GCCAGCTGACTAGAGAGCTTACCACTCCCTGCTCCCATTCGCGCGGAGACCGGA 4198
Qy 1585 GAGCBAATAGAACCTATTATGCAAG--GCAAGAGAGGTAGCGTGAAGCA 1638
Db 4199 GTGCTCAGCTGCGCTTGAAGAACTAGATATACAGCAGGTGAGGGGTGTGAACA 4258
Qy 1639 CCACGAGACATACAGGTTGACAAATTAACCAAGCGCAAGAGAAATTTGGCTTTACGCT 1698
Db 4259 CTTCGACGCGGTGAAGATCACCGCACAGCCAGACGATCTACTAGAAATTAAGTA 4318
Qy 1699 ATACTTTACCCAGCGGATTTGAATATGAAAACTGCGTGTATCCACCATTTGGCG 1758
Db 4319 GTTCTGTCCCGAGACCGTCTCAAGACTCCAAAGTTGCCCCCGTGCACCTCTAGCA 4378
Qy 1759 GAACAACTACTGTAATGACTCAAAAGGTAGGCGAGGAGATPACAAAGTGCAGCCATAC 1818
Db 4379 GAGAGGTGAAATTAATTAACATTAACGAGAGGCGCGGTTTACAGGTGACGGATAT 4438
Qy 1819 CACGTTAAGTCAATTTGACCAAGAAAGGACGCGTCCCTGTTCAAGACTTCCAGGCAAT 1878
Db 4439 GACGCGAGGCTCTACTACATGATGATTCGACATTCGGTCCCTGATTTCAAGCTTTG 4498
Qy 1879 AGTGAAGCGCTACGATGTTTTCACGAGAGGAGTTGCTAAACAGATCTGCACAC 1938
Db 4499 AGCAGAGCGCCACTATGAGTACAAAGAAAGGAGTGTCTCAACAGAAACATATACAT 4558
Qy 1939 ATGCAATCAACGAGAGGCGTAAACATGACGAAAGTATCTTAAGCTGTAAAGACT 1998
Db 4559 ATTGCGTTACGAGCGGTGCTGACCAACGAGAGAACTACGAAAGTACAGCT 4618
Qy 1999 CAGGACACAGACTCAGAAATAGTCTTGCATTTGACGACGAAAGTGTAAAGCGAA 2058
Db 4619 GAAAGAACTGACGCGAGTACGTCTTGCAGTATTAAGAAATGCTGCTCAAGAGAGAG 4678
Qy 2059 GACGAGGTCCTTGTGCTTAACCGGTATCTGTAGATCCACATTTCAAGATTTGCG 2118
Db 4679 GAAACGTCGGGTTGTGTGTGGAGAGCTAAACCAACCCCGTTCATATGATTCGCC 4738
Qy 2119 TACGAGTCTCAAGACAGACGACGACCTCAAGAAATCCCAACCATCGGACTAT 2178
Db 4739 TACGAAAGGCTGAAGATCAGCGCGTGGCACATATAGACTACAGATAGAGGCTTT 4798
Qy 2179 GAGTGTCCAGTTCAAGTAAATCTGAATCATCAAAAGCGCTGTGACTAAGAAAGATCTG 2238
Db 4799 GGGGTTCCGGATACGCAAGTCTGTATATTAAGAGCTCTGTGACCAACAGATCTG 4858
Qy 2239 GTTGTAGTGCAGAAAGAAACCTGCGCAGAAATCATCAGGATGTAAAGAGATAGA 2298
Db 4859 GTACACAGCGCAAGAGAGAACTGCAAGAAATATGTTAACGCTGAAGAAACACCGC 4918
Qy 2299 CGTATGATGTGTCTAGAGCTGTGATTCAGTGTCTTAAATGGGGTTAAAGCACCCC 2358
Db 4919 GGGAAAGGGAACAAATAGGAAACAGTACTCATCTGTCAAACGGGTGTCTGTGCC 4978
Qy 2359 GTTAACACTGTATCATTTGATGAGGCAATTTGCTGCATCAGGAGAGGCTGTGGCACTG 2418
Db 4979 GTGACATCTCATATGTGAGAGAGCTTTGCTTGCAATTCGGTATCTGTGCTGCGCTA 5038
Qy 2419 ATTGCATCTGCAACT--AAGAAAGTGTATTTGTCGGGAGACCCAAACATGCGGC 2475
Db 5039 ATTGCTCTGTAAACCTCGAGCAAGTGTGTATGCGAGACCCCAAGCAATGCGGA 5098

QY	2476	TTCTTTAACAATGATGTGCTGAAAGTAACATTTTAACTGACATATGACCTGAAAGTGCAC	25353
Db	5099	TTCTTCAATATGATGACAGCTTAAAGGTGAATCTTCAACCAACATCTGCACTGAAGTATGT	5158
QY	2536	CATAAACAATCTCTAGAGAGGGGACACAGACCTGTAACGGGCATGTCTCCACGCTCTTC	25955
Db	5159	CATTAATAATATATCCAGACGTTTGACGGGCTCCAGTCAGCGGCATGTGTCTAACGTTGCAC	5218
QY	2596	TACGACAAAGCAATGAAACGAGTTAAACCATGTGCTGATTAATAATCATATGATATCCACA	26555
Db	5219	TACGGAGGCAGATATGCGACGACCAACCCGTGCAACAACCATATATATGACACACACA	5278
QY	2656	GGGACCAACAAGCCGCAACAAGATATCTGATTTCTAACCTGTTTCAAGATGGGTGAAA	2715
Db	5279	GGACAGACCAAGCCCAACCCAGAGACATCGTTGTAATCATGCTTCCAGGCTGGGGCAAG	5338
QY	2716	CAGCTACAGATTTGACTACAAAATACGAATATCAGTATGGGGCTGCAATGCCAAGACTT	2775
Db	5339	CAGCTGCAAGTTGAGACTTACCGTGAGACCAAGATCAGACAGACGACATCTCAAGGCTTC	5398
QY	2776	ACGCGGAAGGCGCTTATATGCTGTCAAGTACAAAGTCAACGAAATCCACTCTACGCGAG	2895
Db	5399	ACCCGCAAGGGGTATACGCCGTAAAGCAGAAAGGTGAATGAATAATCCTTGTATGCCCT	5458
QY	2836	ACTTCTGAGCAGTGAACGTGTTACTTACACGCAACGAAAAACGATTTGTCTGAAAGACG	28959
Db	5459	CGCTCGAGGACAGCTGAATGTACTGCTGACGCGCACTGAGATGATGGCTGTGTGAAAAACG	5518
QY	2896	CTAGCTGTGATCCCTGATTAAGACACTTACAGTAAATATCCGGGGATTTCAAGGCT	29555
Db	5519	CTGGCGGGCGATCCCTGTGATTAAGGTCTTATCAAACTTCCACAGGGGTAACTTTAACGGCC	5578
QY	2956	TCATTGACGACCTGGCAGCGGAACACAGACGCGAATATGAGCAGCGGTTCTTGATTAAGCG	3015
Db	5579	ACATTGGAAGATGGCAAGAAACACACAAATATGAAGGTGATGAAGACCGGCT	5638
QY	3016	CAGACAGCTGATGTGTTCCAGAAATTAAGGTGAACGTCTGCTGGGCGAAGGCTTTAGAGCA	30759
Db	5639	GGCGCTGTGAGCGGGTTCCAGAACAAACGAAACGATGTGTGGCGMAAAGCTGTGCTCT	5698
QY	3076	GTCCTGGCCACGCGCCACATTTGTGCTGACGAGACAGCATGTGGAGAC--GTGCAACCA	3132
Db	5699	GTCCTGGCACCTGGCCGGAATCAGATTGACACAGAGGAGTGGAGCACCATATTAACAGA	5758
QY	3133	TTCAAGCATGACAGAGCGTACTCAACCTGAAATGAGCATGAACTTTTGACACAGATTC	3192
Db	5759	TTTAAGGAGGACAGAGCTTACTCTCAATGTGGGTCTTGAAATTTGGACCAAGTAC	5818
QY	3193	TTTGGAGTAGACCTGGAACAGTGGGTTATTTTCCGCTCTACCGTGCACCTACTTACACAG	3252
Db	5819	TATGAGATGTAACCTGGAACAGTGGGCTGTTTCTGCCCCGAAGGTGTCCCTTATTTAGAG	5878
QY	3253	GATCAGCACTGGGATTAATCTGCGCCAGGGAAGAACTGATGGGCTTAATATGAGAGTGA	3312
Db	5879	AACAACCACTGGGATTAACAGACCTGTGGAAGATGTATGATTCAATGCCGAACAGCT	5938
QY	3313	AAGGATGTGTACCGCCGATATCCGTGATCTCAAAAAGCGGTTGACACAGGAGGTAGCT	3372
Db	5939	GCCAGGCTGGAAGCTGACATATCTCTCTGAAGGGGCGGTGCAATCCGGGCAACAGGCA	5998
QY	3373	GATATTAAGGAATTAATACCATCAAGGACTACTCTCCAACTTAATATGTGCTTCCATTAAT	3432
Db	5999	GTATATCGAGAAAGAAAATCCAAACGCTTTCTGTGCTGACAAATGTATATCTTATCAAC	6058
QY	3433	CGCCGGTTGCCCACTCGTGTGATCGTTGACCAACAAGGACAGGGGTACAACTGATACAGC	3492
Db	6059	CGCAGGCTGCGCAGCCCTGTGGCTGATGATCAAGACGTTTAAAGGAGTAAAGGTTGAG	6118
QY	3493	GGATTTCCATCTAATAGATGAAGGCAAAATCTGTGTTGGATCGCGCATCTTATACGATT	3552
Db	6119	TGGCTGTGCAATTAAGTAAAGAGGATACACAGTCTGCTGCTGGAGTGAATCAACCTGGCT	6178

QY	3553	CCAGGGAAGAAAGTAGAATCCATGGGTCCATTGCCACTAATAC-----CATCAG	3603
Db	6179	TTGCCTGACGCAAGGCTCATTGGTGTTCACCGCTGAATGTCCAGCGCCGATAGGTG	6238
QY	3604	TGTGATCTCGAATTTGGGAATACCTTAGCATGTCCGGTAATATGACATTATCTTGTCAAT	3663
Db	6239	TACGACCTTAAGTTTAGAATCGTCCGGCTGACGCCCGGAGTTCCGACTTGTGTGTGAAC	6299
QY	3664	GTTAGAGACCCCGTACAGGAACTATCTACCAAGTGCAGAGATCACGCTATCCACAC	3723
Db	6299	ATTCCACGGAATTCAGAAATCCACCACTACAGCAGTGTGTGCACACGCGCATGAACTGT	6358
QY	3724	AGCATGCTTAAGGTGTGAAGGTGTCCACCACTTGAACTATGGCCGGAACATGTGTGCTATA	3788
Db	6359	CAGATGCTTGGGGAGATGCGCTACGATGTGTAAACCCGGCGCATTT--TGATGAGA	6415
QY	3784	GGGTATGGGCTTGGCTGATCGCGCAACGGAATATCATCTGCGGTGGACGGCTCATTT	3843
Db	6416	GCTTAGGATACGCCGATTAATACGCAACCGCTTGTCTCTCTTAAGCAGAAAGTTCC	6475
QY	3844	AGGTTTACCCTGTCTGTCAAGCTTAAGAACTGCGCGAAATACTAGAGTTCTCTCGTG	3903
Db	6476	TCGTCTGCAAGATGTTTGGCGCCGGAATTGTGTCAACGACCAATACGAAGTGTCTTGCTG	6539
QY	3904	TTCTTTGGCAAGACACGACCAACACACATGACAGGACAGACTCGGTGTATGTCTT	3963
Db	6536	TTCTTCCACTTTTGACACGGAAGAGACCTCTACGCTACACAGATGMAATCCAAAGTGT	6599
QY	3964	GACAAACATCTATCAAGGGTCAACAGAGTAGAGGACAGGAGAAAGTCCAGCGTACAGGTG	4023
Db	6596	AGTGGCGGTATATGCCGAGAAACCATGACACGCGCGGTGTGACCATCTTACAGAGTT	6655
QY	4024	ATCAGAGGTGACATTAGCAAGAGCGCTGACCAAGCTATCTTTAATGTCTAATATGCAAA	4083
Db	6656	AAGAGAGCAGACATAGCCACCGTGACAGAAACCGGCTGTGTAAAGCAGTAAAGCCCGCT	6715
QY	4084	GGTCAACCAAGTTCCGAGTGTGCGGTGCTACTGTACCGAAATAGCGCGGTCTTTGAT	4143
Db	6716	GGAACTGTAGGGGATGGCGTATGCAAGGCGCGTGCGAAAGAAATGGCGTCAAGCTTTAAG	6775
QY	4144	AGACAGCCAAATAGCTGTCCGGGACGGCTAGATTGTGAAGCAACGAACCG--CTCATCTATA	4200
Db	6776	GGAGCAGAACACAGATGGGACCAATTAAACATCATATGTGCGCTCGTACCCGTCATTC	6835
QY	4201	CATGCTGTAGACCAATTTTCTTAAGTGTGCGGAAACCGGAGGGCGCACTTAACTCGCA	4260
Db	6836	CACGCTGTAGGGCTTAATTTCTCTGCAACGACTGAAGGGAAGGGGACCGGCAATTGGCC	6895
QY	4261	GCTGCTTACATGACATAGCGTTCATCGTCAACGCTGACGGATTAACAAAATATCACTA	4320
Db	6896	GCTGCTTACCGGGCAGTGGCGCGGAGTAAGACAGTGTCACTGAGAGCGTGAACATC	6955
QY	4321	CCGCTACTGTGTAACCGGATCTATTTCTGTGTGGCAAAAGTGAAGTATCATTTGAT	4380
Db	6956	CCGCTGCTGTCAACAGAGATGTTCAGCGGCGGGAAGATAGGTGCAAGCAATCCCTCAAC	7015
QY	4381	CACCTGTTCACTGCTTTGCACTACAGGATTCGCGATGTCACTATATATTTGCTGGATAAA	4440
Db	7016	CATCTATTACAGCAATGAGCGCACCGACGCTTACTGTACACATCTACTGACAGACAAA	7075
QY	4441	CAATGGAGACACAGGATTAATGAGGCAATTCACCGCAAAAGAAACGTCGAATTTCTGAT	4500
Db	7076	AGTTGGGAGGAAGAAATCCAGGAAGCCATTGACATGAGAGCGGCTGTGGAAATGCTCAAT	7135
QY	4501	GATGACAAAGCAATGAGACATTGATTTGGTCAAGGTTCAACCCAAACAGTCTTTTGGCAGGC	4560
Db	7136	GATGACGTGAGCTGACCAACAGATTGGTGAATGACCCCGGACACACAGCTGTGGGT	7195
QY	4561	AGACCAAGTTACTCCGTCAATGAGGGAAGTTGATTCATATCTCGAAGGATACGATTC	4620
Db	7196	CGTAAAGGCTTACAGTACCATGACGCGGTGCTGTACTGTCTTTGAAGGTGCAAAATTC	7255
QY	4621	CATAGACCCGCAAGACATTTGCCGAATCATATGCAATGTGGCCCAACAAATCTGAGCT	4680

Db 7256 AACGAGCTGCTATGATATGATGAGATGATGAGATGCTGATGCTGAGCCCAAGCTGCAAGAGCA 7315
Qy 4681 AATAGACAGATTTGCTGTTATCATCTCTGGGAGAGATATGTCACAGTCTCCCTCCAAATGC 4740
Db 7316 AACCAAGAGATATCTATATGAGCTGAGCCAAACAAATGACAACTATGATCCAAATGT 7375
Qy 4741 CCATAGAGAGATGAGAGCGCTGCTCCTCCTCAGACACTTCCATGCTGTATATAC 4800
Db 7376 CCGGTGAACATTTCCGATTTATCAACACTCCAGAGAGTCCCTGCTGTGCGCTAC 7435
Qy 4801 GCTATGACGCTGAGCGGATATACAGTTGCGCTCTGCAAGAAAGAAACAGTTGCGCTA 4860
Db 7436 GCAATGACAGACAGAGATGCGCGCTTATGATCACAACAAATTAAGATGATGCTT 7495
Qy 4861 TGCCTATCATCTCTGTTGCTCCCAATGATCAGATCAGAGCGCTGCAAGAACTACATGACG 4920
Db 7496 TGCCTATCTTTTCCCTCCCAATATCAGATGATGAGGATGAGAAAGTAAAGTGGAG 7555
Qy 4921 AACCAAGTCTGTTTTCAGGCGGTATACACCGGCTGTACACCCGAGAAATGACGCGAA 4980
Db 7556 AAGTCTCTCTGTTGACCGGATGATCTTCAATGATGATGCTGCGGAGATATGCGCA 7615

RESULT 9

US-09-190-246-4
; Sequence 4, Application US/09190246
; Publication No. US20030180257A1
; GENERAL INFORMATION:
; APPLICANT: Partridge, Mark
; APPLICANT: Li, Xiaomao
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: ALPHAVIRUS VECTORS FOR PARAMYXOVIRUS VACCINES
; CURRENT APPLICATION NUMBER: US/09/190,246
; CURRENT FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 8100
; TYPE: DNA
; ORGANISM: Semliki Forest virus
US-09-190-246-4

Query March 29.5%; Score 1473.8; DB 12; Length 8100;
Best Local Similarity 57.9%; Pred. No. 0;
Matches 2779; Conservative 0; Mismatches 1982; Indels 36; Gaps 8;

Qy 214 ATTGGAAGTGGCCGCTGAGACATGACATTCATTCACCGCTATCTATCTGCTCT 273
Db 1 ATCGGAGTGGCGCTTCCAGAGAAATGATGCTACGCAAAATCACTGCTATGCTCT 60
Qy 274 ATGTAAGCGCTGAAGACCGGACAGACTACAGGATGACAGAAAGACTTAAGAAA-- 331
Db 61 ATGGCAGCGCAGAGAACCCCGAAAGGCTCGATGCTACGCAAGAAACTGGCAGCGCC 120
Qy 332 ----GTGACATTACCGAACAAGATATGCTTAAGGCGGAGACCTGCTGAGAGTATG 387
Db 121 TCCGGAGAGGTGCTGATAGAGATGCGAGAGAAATACCGACCTGCAACCGTATG 180
Qy 388 TCAACACAGACGAGAGACTCATCTCTGTATGACACAGAGCGCACTGTAGTAC 447
Db 181 GCTACGCGCAGACGCTGATATCTTACCTTTGCTGCTACACAGAGCTGCTGTACG 240
Qy 448 TTGGAAGTGAAGATATCAAGATGTGTAGCGACGTCACCGACCGACATCATCTAC 507
Db 241 GCAACCGAAGTGGCGCTATACAGAGACGTGTATGCTGTACACCAACATCGCTGAC 300
Qy 508 CACGAGCGCTTAAGAGATTAAGACATTTACTGATAGGCTTTGACACGCCCTTTT 567
Db 301 CATGAGCGATGAAGAGGTGTGAAACGCGCTATGATGAGTTGACACCAACCCGCTT 360
Qy 568 ATGTACAAAAACATGACGAGTTCTTACCTACTTACAAACGAACTGGGCTGACGAGAG 627

Db 361 ATGTTTACCGCGCTAGACGCGGCTATCCAACTAGCCACAAACTGCGCGACGAGCAG 420
Qy 628 GTATTTGAACAGCTATACATTTGGCTCTGCTGATCTCAGATCTTACAGAGACAGCTTGA 687
Db 421 GTGTTAAGCGCCAGAGAACTATGAGCTGTGACAGATCTTGACTGAGGAGAGCTGGC 480
Qy 688 AAATCTCAATCTTTAGAGAGAGGCTCAACTCTATTAATAGATCATATTCTCGTT 747
Db 481 AAATCTCAATCTTCCGCAAGAGCAATTGAATCTTGCACACAGTCACTTCTCGTTA 540
Qy 748 GATTCAACATCTACAGAGAAATGATATCACTGTATGATGCTGCGACTTCCAACTG 807
Db 541 GATATCTATTTGATACATGAGACAGAACTCTAGAGAGCTGCGACTTACCTCCGTA 600
Qy 808 TTCCACTTGAAGAGAAAGTAACTTACAGAGTATGATGAGACATTTGCTAGCTGAA 867
Db 601 TTCCACTTGAAGAGTAAACAAATCTTATGAGTATGATGATGATGATGATGATGAA 660
Qy 868 GGGTACGTCATCAAAAAGATATGATGATGATGATGATGATGATGATGATGATGATG 927
Db 661 GGGTACGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Qy 928 GCGTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 987
Db 721 GCGTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Qy 988 GAGAGGTTCTTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1047
Db 781 GAAAGATCTCATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Qy 1048 GGGATTTGCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1107
Db 841 GCGATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
Qy 1108 CAAAGATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1167
Db 901 CAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
Qy 1168 TTACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1227
Db 961 CTTCGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
Qy 1228 GACGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1287
Db 1021 GATGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Qy 1288 AAGACCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1347
Db 1081 AAAACGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
Qy 1348 CTGCGCTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1407
Db 1141 CTTCAAGATTTAATCTGTTGATGATGATGATGATGATGATGATGATGATGATG 1200
Qy 1408 TTCCGCGTGGCTCAAGCTGCTGCTTGAACCAACTGTCAACCCGACCGGCTATTACA 1467
Db 1201 GTCAATCAGCATTAAGATGCTTTTGGCCAAAGAGCAAGAGAGAG--GTTAATCT 1257
Qy 1468 ATGGCCGATGAGAGATCTGCTGCTTACAGCAAGAGCTGAAGAGTGGCTGACGG 1527
Db 1258 GTTCTGACGCGTGTGATGAGCAAGAGATGCTGAACAAAGAGAGAGAGAGTGGAGGCC 1317
Qy 1528 GAAGAGATCAGAGAGCGCTGCACTCTTCTCTCTGAAATATAGAAAAGAGACCTGAG 1587
Db 1318 GAGCTGATGAGAGAGCTTACCAACCTCTGCTCCCATGCGCGCGGAGAGAGAGGTC 1377
Qy 1588 GCAAGATGAGACCTTATGAGAGAG-----GACAGAGCAGAGTGGAGAGACCA 1641
Db 1378 GTGACGTCGATGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1437
Qy 1642 CGAGACATCATGAGGTGACAAAGTTACCGAGCGAAGAGAAAGATTTGGGCTTACGCTATA 1701

1438 GCAGCGCGTTGAAAGTACCGCAGCCGAAACGATCTAGAAATTACCTAGTT 1497
Qy 1702 CTTTACCCCGAGCGGATTTGAAATAGTAAAAATCGCGTATCCACCATTGGCGGA 1761
Db 1498 CTGTCCCGGACCGCTGCTCAAGAGCTCAAGTTGGCCCCGTCACCTCTTACAGAG 1557
Qy 1762 CAAGTACTGTTAATGACTCAACAAGGTAGGGCAGGAGATACAAAGTCCAGCCATACAC 1821
Db 1558 CAGGTAAATATATATACATTAACGAGAGGGCGCGGTTACCAAGTGAAGATATAC 1617
Qy 1822 GGTATGATCTATTTGATCCAGAAAGGACGGCGGTCCTGTTCAAGACTTCCAGGCAATTGAGT 1881
Db 1618 GCGAGGGCTCTTACATCAATGTGGATCGGCATTCGGTCCCTGAGTTTCAAGCTTTGAGC 1677
Qy 1882 GAGAGCGCTAGCATCGTTTTCAGAGAGGGAGTTGTTAAACAGATACCTGACACATC 1941
Db 1678 GAGAGCGCCACTATGTGTATCAACGAAAGGAGTTGTCACACGAAACTTATCCATATT 1737
Qy 1942 GCAATCAACGAGAGCGCTAAACATGACGAGATCTATTAAGACTGTAAAGACTGAG 2001
Db 1738 GCCGTTCACGAGCGCTGCTGAACACGACGAGAGAACTACGAGAAAGTCAAGACTGAA 1797
Qy 2002 GACACAGACTAGAAATACGTCTTGCATATTGACGACGAAAGTGTATTAGCGAGAGAC 2061
Db 1798 AGAATGACGCGGATCGTGTTCGACGTAGTAAATAATGCTGCGTCAAGAGAGAGAA 1857
Qy 2062 GCAGGTCCCTTGTGCTAACCGGTGATCTGTGATTCACCATTTACAGAGTTGCGTAC 2121
Db 1858 GCGTGGGTTGGTGTGTGTGGAGAGCTAACCAACCCCGTTCATAGAAATTCGCTTAC 1917
Qy 2122 GAGAGTCTCAAGACACGACGACGACCTCAACAAGTCCCAACCATCGAGCTATAGGA 2181
Db 1918 GAAAGGCTGAAGATACAGCCCGTCGSCACCATTAATAGACTACAGTAGTAGAGTCTTTGG 1977
Qy 2182 GTGCCAGTTCAAGTAAATCTGGAATCATCAAAAGCGCTGTGACTAAGAAAGATCTGTT 2241
Db 1978 GTTCCGGATACGAGCAGTCTGTATTATTAAGACCTCGTGAACCAACGATCTGCTC 2037
Qy 2242 GTGAGTGGAGAGAGAAACCTGCGCAGAAATCATCAGGAGTGTAGAGAGATGAGAGCT 2301
Db 2038 ACCAGCGGCAAGAGAGAACTGCCAGAAATATGTTAACAGGTGAACAGACCGCGG 2097
Qy 2302 ATGATGTTGCTGCTAGGACTGTGCATTCAGTGTCTTCAAAATGGGTTAAACACCCGCTT 2361
Db 2098 AAGGGGCAAGTAAAGGAAACAGATGATCCATCCGCTAAACGGGTGCTGTGCTG 2157
Qy 2362 AACACTCTGATCATTTGATGAGGCAATTTGCTGCATGACGAGGACGCTGSCACTGAT 2421
Db 2158 GACATCTATATGTGACGAGGCTTTCGTTGCCATTCGCGTATCTGCTGGCCCTAAT 2217
Qy 2422 GCCATCGTCAAACT--AAGAAAGTGTATTTGTGCGGGACCCAAACATGCGGCTTC 2478
Db 2218 GCTCTGTTAAACCTCGGAGAAAGTGTATGTACGAGACCCCAACCAATGCGGATTC 2277
Qy 2479 TTTAATCATGATGTGCTGAAGTACATTTTAACCTGACATATGATGATGATGATGAT 2538
Db 2278 TTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2337
Qy 2539 AAAAGCATCTTGAAGAGTGCACACAGCTGTAAACCGGCATGCTCCACGCTCTTTCAC 2598
Db 2238 AAAAGTATATCAAGCGTTGACGCGCTCAAGTCAAGGCGCATGTGTCTACGTTGACATC 2397
Qy 2599 GACACGGAATGAAAGCGGTTAACCCATGTGCTGATTAATATCATATGATATCCACAGGG 2658
Db 2398 GGAAGCAAGATGCGACACCAACCCGTCGAAACAAACCATATATCATGACACACAGGA 2457
Qy 2659 ACCACAAAGCGGCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2718
Db 2458 CAGACCAAGCGGCAAGCGGAGACATGCTGTTACATCTTCCAGGCTGCGGCAAGAGAG 2517
Qy 2719 CTACAGATTTGATCAAAATCAAGAAATCATGATGCGGCTGATCGCAAGAGATTACG 2778
Db 2518 CTGAGTGTGACTACCGTGACACGAAAGTCAATGACAGCAGCATCTCAGGGGCTCAC 2577

Qy 2779 CGAAAGGCGTTTATGCTGTGAGTACAAAGTCAACGAAATCCACTACTCCGAGACT 2838
Db 2578 CGCAAGGGGATATACCGCTTAAGCAGAGAGTGAATGAATTCCTTTATGCTCCGCG 2637
Qy 2839 TCTGACAGCTGAACGTGTATTCTTACACGACAGAAAAACCATTTGCTGAGAGAGCTTA 2898
Db 2638 TCGAGCAGCTGAATGTATCTGCTGACGCGCATGAGATAGAGCTGTGTGGAAAAAGCTG 2697
Qy 2899 GCTGTGATCCCTGATTAACACACTTACACTTAATATCCCGGGATTTACCGGCTTCA 2958
Db 2698 GCCGCGATCCCTGATTAAGTCTTACAAACATTCACAGGGGTATCTTTACCGCCACA 2757
Qy 2959 TTGACGATGCGACCGCGAACAGAGCGCATTTATGACCGGTTCTTATTAAGCGGAG 3018
Db 2758 TTGAAAGATGCGCAAGAAACAGACCAATTAATGAAGTGTATGAAGACCGGCTCG 2817
Qy 3019 AACAGTATGTGTTCCAGAAATTAAGTGAACGTCTGCTGGCGAAGCTTTAGAGCAGCTC 3078
Db 2818 CCTGTGACCGGTTCCAGAAACAAAGCGAAGCTGTGTGGCGAAAGCTGTGCTGTG 2877
Qy 3079 TTGCGACGCGCAACATTTGCTGACAGACAGCAGTGGAGAC--GTGACCCATTC 3135
Db 2878 CTGACACTGCGGAAATCAGATTGACAGCAGAGAGTGAACACCATTAATTAACAGATT 2937
Qy 3136 AAGCATGACAGAGGATCTCACCTGAAATGSCACTGAACCTTTTGACACAGGTTCTT 3195
Db 2938 AAGAGAGCAGAGTTACTCTCCAGTGGTGCCTTGAATGAATTTGACCAAGACTAT 2997
Qy 3196 GAGTATGACCTGACAGTGGGTTATTTTCCCTCTTACCGTGCAGCTTACTTACAGGAT 3255
Db 2998 GAGATTGACTGACAGTGGGCTGTTTCTGCCCCGAAAGTGTCTCTATTAAGAGAAC 3057
Qy 3256 CAGCACTGGGATTAATCGCCAGGAAAGACATGTATGGCTTAATAGAGATGACAAAG 3315
Db 3058 AACCCTGGGATTAACAGACCTGTGAGAAAGTGTATGATTAATGACGCAACAGCTCC 3117
Qy 3316 GAGTTGACAGGCGATATCCGTGATCAACAAACGGTTGACAGAGGAGGTGCTGAT 3375
Db 3118 AGGCTGAGAGCTAACAATCTTCTGAGAGGCGAGTGGCATAGCGGCAAGCGCAATT 3177
Qy 3376 ATAAAGATATATACATCAAGAGACTACTCTCAACAAATTAATGTGTTCCATTAATGCG 3435
Db 3178 ATGCGAAGAAAGAAATATCCACCGCTTTCTGTGTCGACAAATGATATTCATCAACGC 3237
Qy 3436 CGGTTGCCCACTGCTGATGCTGTGACCAAGAGACAGGTTACACTGATCAGCGGA 3495
Db 3238 AGGCTGCCGACGCGCTGTGCTGAGTAAACAGACGTTAAAGCAGTAAAGGTTGAGTGG 3297
Qy 3496 TTCTATCTAATGAAGAGGCAAAATCTGTGTGTGATTCGCGCATCTATCAGATTCCA 3555
Db 3298 CTGTCAATTAAGTAAAGAGGTATCCACGCTCTGTGTGATGATGATCAACCTGAGCTT 3357
Qy 3556 GGAAGAAAGTAAAGTCAATGAGTCAATGCGCTCAATAAC-----CATCAGGTGT 3606
Db 3358 CTTGAGAGCAGGCTCATTTGTTGTACCGCTGAATGTACAGCGCCGAGATAGGTGCTAC 3417
Qy 3607 GATTCGATTTGGGAATACCTAGCCATGTGCGTAAATATGACATTAATCTTTGCAATGT 3666
Db 3418 GACTTAAGTTTAAAGACTGCGGCTGACGCGGCGAGTTGCACTTGTGTTGTAACATT 3477
Qy 3667 AGAGACCCCGTACAGGAACCATCATCAACAGTGCAGAGATACGCTATCCACACAGC 3726
Db 3478 CACACGAAATTAACAATTCACATTAACAGAGTGTGTGACACGCGCATTAACCTGAG 3537
Qy 3727 ATGCTAAGTGTAAAGCTGTTCACACCTGAAACCTGAGCGGAACATGTTGCTATAGG 3786
Db 3538 ATGCTTGGGGAATGTGCTGACAGCTGTAAACCCGCGCATCT--TGATAGAGCT 3594
Qy 3787 TATGGCTTGTGATTCGCGCAACCGAGAAATATCATCTGCGGTGACGCTCATTTAGG 3846
Db 3595 TACGATACCGCGATTAATATCAGGAAGCGGTTTCTCTTAAACAGAAAGTTCTCG 3654

3847 TTATCCGCTGTCTGTACAGCTTAAGACATGCGGAAATATGAGTTCTTCTGTCTC 3906
3655 TCTGCAAGAGTGTGGCCCGGATTTGTGTACACGACATATGAGAGTTCTTCTGTCTC 3714
3907 TTGGGCAAGACAAAGGCAACACACATGACAGAGTCTGGTGTAGTCTTGC 3966
3715 TCCAACTTTTGCACGGAAGAGACCTCTACGCTTACACAGATGAATACCACTAGT 3774
3967 AACATATTAAGGCTACACAGGATGAGGAGAGTCCAGGATGACAGTATC 4026
3775 GCGTGTATGCCGAGAGAGCCATGACACGCGCGGTGTGACATCTTACAGGTTAG 3834
4027 AGAGTGAATTAAGACAGCGCTGACCAAGCTATGCTTAATGCTTAATGCAAGGT 4086
3835 AGAGCAACATAGCAAGTGTGACAGAGCGCTGTGGTTAAGAGCTTAAGCGCCGTGGA 3894
4087 CAACAGGTTCCGAGTGTGCGGTGACCTGTACCGAAATGCGCGCTGTTGTATGA 4146
3895 ACTGTAGGGAGTGTGATGACAGGCGGTGCGAAAGAAATGCGCTTAAAGGGA 3954
4147 CAGCATTAGCTGTGCGGAGCGGTGACCTTGTGAGACGACGCG---CTCATATCAT 4203
3955 GCAGCAACACAGTGGGCAATTAACATTAACATTAATGCGCTGTACCCGTATCAC 4014
4204 GCTGTAGACCAATTTTCTTAAGATGCGGAGACCGAGGCGACCTTAAGCTGACGT 4263
4015 GCTGTAGCGCTATTTCTGTGACAGCTGTAGCGGAGGAGGAGCGGAAATGGCGCT 4074
4264 GCTTACATGACATAGCTGTATGTCAACGCTGACGCGGATTAACAAATATCAATCCG 4323
4075 GTCTACCGGAGTGTGCGCGCGGAGTAAACAGACTGTACAGACAGGCTGACATCCG 4134
4324 CTATCTCAACCGGATCTATTCTGTGGGCAAGATGAGATGACATCATTCATCATC 4383
4135 CTGCTGTACAGAGAGTGTAGCGGCGAGAGATGAGTGTGACGCAATTCCTCATCAT 4194
4384 CTGTCTCTGCTTTCGACACTACGAGTGTGCGGATGTACCATATATTTGTTGATTAACAA 4443
4195 CTATTCACAGCAATGAGACGCGACGCGAGCTGACGATCATCTGTGAGAGCAAAAGT 4254
4444 TGGGAGACAGGATTAATGAGGCAATTCACCGCAAGAAAGCGTGAATTTCTGATGAT 4503
4255 TGGGAGAGAAATTCAGAGAGCCATTCATGACATGAGAGCGGCTGTGAGTTCTCAATGAT 4314
4504 GACAAGCAGTGAATGATGATGAGGCAAGTTGATTCATCTGAGAGTACAGATTCAT 4623
4315 GACGTGAGCTGACACAGACTTGTGTGAGTGTGACCGGACAGCGCTGTGGGTGT 4374
4564 CCAAGTTACTCTGCTCAATGAGGCAAGTTGATTCATCTGAGAGTACAGATTCAT 4623
4375 AAGGCTTACAGTACACAGTGTGAGTGTGATCTGTGATCTTGAAGTACGAAATTCAC 4434
4624 CAGACCGCAAGACATTTGCGGAATTCATGCAATGTGCGCCAAACAAATCTGAGGCTAT 4683
4435 CAGCTGTATTTGATGTGAGAGATGACTGACGTTGTGGCCAGACTGCAAGAGGCAAC 4494
4684 GAGAGATTTGCTGTACATCTGTGGGAGAGTATGTCAGATTCCTGCAATGCGCA 4743
4495 GAACAGATATGCTATACGCGCTGGCGGAACAAATGACAAATGATTCAGATTCG 4554
4744 GTAGAGAGTACAGAGGCTGTGCTCACTGACCTGACACTTCATGCTGTATTAATTCCT 4803
4555 GTGAACGATTTCCGATTTCAACACTTCCAGAGAGTGTGCTGTGTGCGCGTACCA 4614
4804 ATGACGCTGAGCGGATTAACAGTTGCGCTGTGCGAAGAAAGCACTTGTGCGGTATGC 4863
4615 ATGACAGAGAGAGGATGCGCGCTTGTGTGCTGACACCAAGTTAAAGATGTGTGTGC 4674
4864 TCATCATTTCTGTTGCGGAATGACAGATTCACAGCGCTGCAAGAGTCACTGACAGCAA 4923
4675 TCATCTTTTCTGCTGCGGAATATCATGTAGATGGGTGACAGAGTAAATGTGCGAGAG 4734
4924 CCAATCTCTGTTTACAGGCGTGTACACCGGCTGTACACCCAGGAATGACGCGGA 4980

Db 4735 GTTCTCTGTTTCAACCGGACGTAACCTTACAGTGTATGTCGGGAGATATGCGCA 4791
RESULT 10
US-09-507-362-102
; Sequence 102, Application US/09507362
; Publication No. US20030096397A1
GENERAL INFORMATION:
APPLICANT: Dubensky Jr., Thomas W.
; Polo, John M.
; Belli, Barbara A.
; Schlesinger, Sondra
; Dryga, Sergey A.
; Frolov, Ilya
TITLE OF INVENTION: RECOMBINANT ALPHAVIRUS-BASED VECTORS
WITH REDUCED INHIBITION OF CELLULAR MACRO-MOLECULAR
SYNTHESIS
NUMBER OF SEQUENCES: 125
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed Intellectual Property Law Group PLLC
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/507,362
FILING DATE: 18-Feb-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 930049,45706 /1196,011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
LENGTH: 8000 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 102:
US-09-507-362-102
Query Match 26.0%; Score 1300.4; DB 11; Length 8000;
Best Local Similarity 56.0%; Pred. No. 0;
Matches 2819; Conservative 0; Mismatches 2131; Indels 80; Gaps 16;
23 ATATGAAAGAAATTCAGTGTGATGATGCTGACACCGGTATGTCAAGTCTTAAGC 82
64 AGAGCCAGATGATTAAGTGTAGTACGACCCCGAGTCTGTTGTGTGCAATGCAAA 123
83 GAGCGTTTCAAAATTTGAGATGAGCAAGGCAAGGCTGACTGACATGACATGCCAATG 142
124 AAAGCTTCCGCAATTTGAGGTAGTACACAGCAGGTCACTCCAAATGACATCTAATG 183
143 CCAAGCGTTTTCGATGCTGCAACAAGCTCATTTGAGAGGAGAGTGCACCGGAGCAAG 202
184 CCAAGCAATTTTGCATCTGCGCCAGTAACTTAATGAGCTGAGAGTTCTTACCAAGCA 243
203 TTATCTTGACATTTGAGAGTGTGCGCGCTGACAGATGACATTCATTCACCGCTATCAT 262
244 CGATCTTGACATGAGCAGCGCACCGCTGTAGAAATGTTTTCGAGACACAGATCAT 303
263 GTATCTCCCTATGATGAAGCGCTGAGACCGGACAGACTACAAAGGTATGCAAGAAAG 322

Db 304 GTGTCTGCCCATGCGTAGTCCAGAAAGCCCGACCGCATGTGAATATGCCAGTAAC 363
Qy 323 TTAAAGAAA-----GTGACATTACCGACAGAAACATAGCTCTTAAGGCGGAGACTGCG 376
Db 364 TGGCGGAAAACCGTGCAGGATTACAAAGAAAGCTTGATAGAAAGATTAAAGATCTCC 423
Qy 377 TGGAGTCATGTCAACACGACGAGAGACTCATCTGTGTATGACACAGACGCA 436
Db 424 GGACGTCATCTATAGCGCGGATGCTGAACACATCGCTCTGCTTTCAACAGATGTTA 483
Qy 437 CGTGAAGTACTTTGGAAAGTGTAGCATTAACAAAGATGTGTACGAGTCCATGACCGA 496
Db 484 CCTGCAACATGCGTCCGATATATCCGTATGACAGACGCTGTA---TATCAACGCTCCCG 540
Qy 497 CATCAATCTAACACGAGCGCTTAAAGAGTGTAGACAAATTACTGATAGGCTTTGACA 556
Db 541 GAACTATCTATCATAGGCTATGAAAGGCTGCGACCTGTACTGATGATGGCTTCGACA 600
Qy 557 CGACCCCTTTTATGTACAAAACATGCGCAGTTCCTACTACTTACAAACAGAACTGGG 616
Db 601 CCACCCAGTTCATGTCTCGGCTATGCGAGGTTTCGTACCTCGCTGTAACACCAACTGG 660
Qy 617 CTGACGAGAGATTTGGAAGCATTAACATTGGCTCGGTGTACTTGATCTTTCAGAGA 676
Db 661 CCGACGAGAAAGTCTTGAAAGCGGTAAACATCGGACTTTGACGACAAAGCTGAGTGAAG 720
Qy 677 GCAGGCTTGGAAAACCTCAATCTTAGAAGAAAGAGGCTCCAACTACTTAATAGTCA 736
Db 721 GTAGGACAGAAAATTTGATTAATGAGAAAGAGAGTGAAGCCGGGTCGGGGTTT 780
Qy 737 TATTCCTGGTTGGTTCAACAATCTACACAGAAATAGATCACTGTTACGTAGCTGGACATC 796
Db 781 ATTTCTCGTGAGATGTGACACTTATCCAGAACACAGAGCCAGCTTGACAGACTGGCATC 840
Qy 797 TTCAAAGGTGTCCACTTTGAAAGAAAGTCTTAACCTTACAGATGANTGGAACCATTTG 856
Db 841 TTCCATCGGTGTCCACTTGAATGGAAGCACTGTCGACCTTGCGGTGTGATAGATGG 900
Qy 857 TCAGCTGTAAGGATACGTCATCAAAAAGATTAACATCAAGCCGACAGACTATACGTAAG 916
Db 901 TGAGTTGGAGAGGCTACGTAAGTGAAGAAATACACATGCTCCCGGATCACGGAGAAA 960
Qy 917 TTGAGAACTTGCGCTGCACAAATGCATCGCGAGGTTTCTTGAGTTGCAAAAGTCAAGATA 976
Db 961 CCGTGGGATACCGGCTTACACACATAGGAGGCTTCTGCTATGCAAAAGTTACTGACA 1020
Qy 977 CGCTGCGCGGAGAGGGTTCTTTTGTGTGTGTAGATGATGATACAGGACACACTTGG 1036
Db 1021 CAGTAAAGAGAAAGCGGTATCGTTCCTGTGTGACGTAATCCCGGCACCATATGGC 1080
Qy 1037 ATCAATATGACAGGATTTCTGACACTGACGTTAGTGTGATGATGACGACAAAACATAATGG 1096
Db 1081 ATCAATATGACGATGTTAATGACGACGAAATATCACTGACGATGACAAAACCTTCTGG 1140
Qy 1097 TTGGGCTCAACCAAGATTTGTCTCAATGTGTAGACGCAAGAAATACTAACAACATGC 1156
Db 1141 TTGGGCTCAACCGAGATTTGATTTAAGGTAAGACTTAACAGAACACCAACACATATC 1200
Qy 1157 AGAACTATCTATTAACAGTGTGCGCCAGGCGTTTTCAGAGTGGCGCGTGAACATCGTG 1216
Db 1201 AAAATTACTCTTCTCGCATCATAGCACAAAGGTTTCAGAAATGGGCTAAAGAGCCAAAG 1260
Qy 1217 CCGACTTTGACGACGAGAAAGAACTAAGGGTCCGGAGGCACTCTTACTATAGGGTGTCT 1276
Db 1261 ATGATCTTGATTAACAGAAAGTGTGGGTACTAGAGAAAGCAAGCTTAAGTATGCTGCT 1320
Qy 1277 GCTGGGCTTTCAAGACCCAGAAATCACTCCATCTCAAGAAAGCTGGTATGCAAAACA 1336
Db 1321 TGCTGGCGTTTGCGACTAAGAAAGTACATCTGTTTATGCGCCCACTGGAACGACACTT 1380
Qy 1337 TTAAAGAAAGTACCTGCGCTCTTTGACTCATTTGTGTATTCACGCGCTTACAGCCACGG 1396
Db 1381 GCGTAAAGTCCCAAGCTCTTTTAAAGGCTTTTCCATGTGTCGATGAGACGACTCTT 1440

Qy 1397 TCGATATGAGGCTTCCGCGTAGAGCTCAAGCTGCTGTTGAACCA---ACTGTCAAACCG 1453
Db 1441 TGCCCATGTGCTGAGGAGAGAAATTGAATCTGACATTGCAACCAAGAAAGAGAAAAAC 1500
Qy 1454 CACCGGCTATTAACAATGCGCGATGTGAGCATCTGCGTGGCTTACAGCAAGAGCTGAAG 1513
Db 1501 TGCTCAGAGTCTCGGAGGAATTAAGTATGAGAGGCCAAGGCTGCTTTTGAAGATGCTCAG 1560
Qy 1514 AAGTGTGACGCGGAGAAAGATCAGAGAAAGCCCTGCACTTGTCTCCCTGAATATGAAA 1573
Db 1561 AGGAAGCCAGAGCGGAGAAAGCTCCGAGAAAGCACTTCCACATTAATGTGCGACAAAGGCA 1620
Qy 1574 AAGAGACGCTAGAGCA-----GAAAGTACCTCATTAATGCAAGAGCGAGAGAGC 1624
Db 1621 TCGAGGACAGCCGACAGAGTGTCTGCGAAAGTGAAGGGGCTCCAGGCGACATCGAGAGC 1680
Qy 1625 GTAGCGTGAAGACACACAGAGACACATCAGGAGTACAAATTACCAGGCGAAGAGAGA 1684
Db 1681 CATTAAGTTGAACCCCGCGGCTCACGTAAGGATTAATCTCAACCAATATGACCGTATGA 1740
Qy 1685 TTGGCTCTTACGCTATCTTTCACCCACGCGGCTATTGAATGTGAATAAATCTGCGTGT 1744
Db 1741 TCGGACAGTATATCGTGTCTCGCCAACTGTGTGTAAGAAATGCCAAACTCGCACAG 1800
Qy 1745 TCCACCAATGGCGGAAACAAGTACTGTAAATGACTCACAAAGGTAGGGCAGGAGATCA 1804
Db 1801 CGCACCCGCTGACAAATCAGGTTAAATCATTAACAATCCGGAAGATGAGAAAGTACG 1860
Qy 1805 AAGTGAACCAATACCAACGCTAAGGCTATTTGATACAGAAAGGACGCGGCTCCTGTTCAAG 1864
Db 1861 CGGTGAACCAATAGACCGCTAAAGTATGATGCGACAGAGAGGTGCGGATACATGCGCAG 1920
Qy 1865 ACTTCAGGCAATGAGTGAAGCGCTATGATGCTTTTCAAGAGAGGAGTTCGTAACA 1924
Db 1921 AATTCCTGACACTGAGTGAAGCGCACGCTTAAGTGTACAAACGAAAGAGTGTGGAACC 1980
Qy 1925 GATACCTGACCAATCGCAATCAACGAGAGGCGCTAAACATGACGAAAGTACTTA 1984
Db 1981 GCAACCTATACCAATGCTCATGATGAGCCCGCCCAAGAAATACAGAAAGAGAGTACA 2040
Qy 1985 AGACTGTAAAGACTCAGAGACACAGATCAGTAATCGTCTTCGATTAATGACGACAAAGT 2044
Db 2041 AGGTTACAAAGGCAAGCTTGACAGAAACAGAGTACGTGTTGACGTGACAAAGAGCTT 2100
Qy 2045 GTGTTAAGCAGAGAACCGAGGCTCCCTGTGTGCTTAACCGGTGATTTGATGATCCAAAT 2104
Db 2101 GCGTTAAAGAAAGAAAGCTCAGGCTGTGCTCTCTCGGAGAACTGACCAACCTCCCT 2160
Qy 2105 TTCAAGATTTTCCGTAAGAGAGTCTCAAGACACGACGACGACCTCAAAAGTCCAA 2164
Db 2161 ATCATGACCTAGCTCTGAGAGGACTGAAGACCCGACCTGCGGTCCGTAAGAGTCAAAA 2220
Qy 2165 CCATCGAGTCTATGAGTGCAGGTTCAAGTAACTGGAATCATCAAAAGCGCTGTA 2224
Db 2221 CAATAGAGATGATAGGACACCGGCGTGGGGAATCAGCTATTTCAAGTCAACTGTCA 2280
Qy 2225 CTAAAGAAATCTGTTGTGAGTGTGAGAAAGAAAGAACTGCGCAGAAATCATCAGGATG 2284
Db 2281 CGGACAGAGATCTGTTTACAGCGGAGAAAGAAATTTGTGCGAAATTTGAGGCGGACG 2340
Qy 2285 TAAAGAGATGAGAGATGATGATGTTGTGCTGTAAGACCTGATTCAGTCTTCAATG 2344
Db 2341 TGTAAAGACTGAGGGGATGAGATTAAGTGAAGAAAGTGAATTCGTTATGCTCAACG 2400
Qy 2345 GGGTTAAGCACCCTGTTAAACTCTGTATCAATGATGAGGCAATTTGCTGCATGACGAGA 2404
Db 2401 GATGCAAGAAAGCGGTAAAGTGTGATGCTGTAAGAGAGCGTTCGCGTCCAGCGAGAG 2460
Qy 2405 CGCTGCGCACTGATTCATGCTCAACCC---TAAAGAAAGTGTATTGTGCGGGAACC 2461
Db 2461 CACTACTTGCTGATGTTGCTATGTCAGGCGCCGCAAGAAAGTATGATATGTCGAGGACC 2520

QY	2452	CAAAACATGCGGCTTCTTTAAACATGATGTGCTGAAGTACATTTTAAAC-----	2512
Db	2521	CCATGCAATGGGGATTCTTCAACATGAATGCAACTTAAAGTACATTCACCCCTGAA	2580
QY	2513	ATGACATATGCACTGAAGTGTAACCTAAAGATCTCTAGAGGGTGCACAGACTGTAA	2572
Db	2581	AAGACATATGCACCAAGACATTTCTTCAAGTATATTTCCCGGCTTGCAACAGCCAGTTA	2640
QY	2573	CCGCCATCGTCTCAACGCTCTTCTAACAAGCCGATGAAGACGGTTAACCATGTGCTG	2632
Db	2641	CAGCTATTTGTATCGACACTGCATTCAGATGGAAAGATGAAACCAAGAACCCGTGCAAGA	2700
QY	2653	ATTAATATCATCATAGATATACACAGGGACCACAAAGCCGACAAAGATATCTGATTCTTA	2692
Db	2701	AGAAATTTGAATTCATATTTACAGGGGGCCACAAAGCCGAGGATATCATCTCTGA	2760
QY	2693	CCTGTTTACAGAGATGGGGTGAACAGCTACAGATTGACATACAAAAATACAGAAATCATGA	2752
Db	2761	CATGTTTCCGGGTGGGTTAAGCATTTGCATATGCATTTCCGGACATGAAGTAAAGA	2820
QY	2753	CTGCGGCTGCATCGCAGAGACTTACGCGGAAAGCGTTTATGCTGCAGTACAAAGTCA	2812
Db	2821	CAGCCGGGGCCTCACAAAGGGCTAACAGAAAAGAGTATATGCGTCCGGCAAAAAGTCA	2880
QY	2813	ACGAGATATCCACTCTACTCGCAGACTTTCTGACACGTGAAGCTGTACTTACAGCAGAC	2872
Db	2881	ATGAAACCCCACTGTACCGCATCATCATCAGACATGTAAAGCATGTGTGCTCACCCGCACTG	2940
QY	2873	AAAAACGATGTCTGGAAGACCGTAGCGGTGATCCCTGATTAAGACAATTACAGCTA	2932
Db	2941	AGGACAGGCTATGTGGAAGAACTTTCAGGGGCAACCTGGATTBAGCAGCCCACTTACA	3000
QY	2933	AATATCCCGGGATTTTCACGGCTTCATTGAGACGACTGACAGCGCCACACGACGCATTA	2992
Db	3001	TACCTAAAGGAAACTTTCAAGGTACTATTAAGAGACTGGGAAGCTGAACAAAGGAATTA	3060
QY	2993	TGGCAGCGGCTTCTTATTAAGCCGACAGACAGTGTGTTCAGATTAAGTGAACGTCT	3052
Db	3061	TTGCTGCATTAACACAGCCCACTCCCGTGCCAATCCGTTACGCTGCAGAACCAACTTT	3120
QY	3053	GCTGGGGCAAGGCTTTAGAGCCAGCTTGGCCACGGGCCACATTTGTGTGACGAGACAGC	3112
Db	3121	GCTGGCCAAAGACATTTGGAACCGATACTAGCCACGGCGGTATGCTACTTACCGTTGCC	3180
QY	3113	AGTGG---GAGACGTTGACCCCATTTCAAGCATGACAGAGCGTACTCACTGAATGGCAC	3169
Db	3181	AGTGAAGGAACCTGTTCCCAAGTTTGGGAGTAGAACAACAATTCGGCCATTTACGCT	3240
QY	3170	TGAATCTCTTTTGACCAAGGTTCTTTGAGTAGACCTGGAACAGTGGTATTTTCCGCTC	3229
Db	3241	TAGACGTATTTTGCATTAAGTTTTCCGCATGGACTTGACAAAGCCGACGTGTTTCTAAAC	3300
QY	3230	CTACCGTGCACCTTACTTAAGS-----GGATCAGACAGTGGGATA	3268
Db	3301	AGAGATCCCACTAACGTACCAATCCCGGATTCAGCAGAGCGGTAGCTATTTGGGACA	3360
QY	3269	ACTGCCACAGGGAACAATGTATGGGCTTTAATAGAGAGTAGCAAAAGAGTTGTCAACGC	3328
Db	3361	ACAGCCCAAGAACCCGCAAGTATGGGTACGATCAGCCATTCGCCGCAACTCTCCGTA	3420
QY	3359	GATATCCGTGCATCAAAAGCGGTTGACACAGGACGGGTAGCTGATTAAGAAATATA	3388
Db	3421	GATTTCCGGTGTTCAGCTAGC---TGGGAAGGGGACACAACCTTGATTTTGCAGACGGGA	3477
QY	3389	CCATTAAGACATCTCTCCACAAATTAATGTGTTCCTTAAATGCGGTTGCCCACT	3448
Db	3478	GAACCAAGATTAATCTCTGACACATTAACCTGTGTCGGGTGAACGCAATCTTCTCAACG	3537
QY	3449	CGTTATGTTGACACAAAGGACAGGGTACAATGATCAACAGCGGATTTCCATCTTAAGA	3508
Db	3558	CCTTAGTCCCGAGTACAAAGAGAACAAACCGGCGGTCAAAAATTTCTTGAACCACT	3597
QY	3509	TGAAGGCAAACTGTGTTGTGTATC---GGGACTCTATCAGCATTCAGGGAAGAAAG	3565

Db	3558	TCAAACACCTCAGTACTTGTGGATTCAGAGAAAAATTGAAGCTCCCGTAAGAGA	3657
Oy	3566	TAGAGTCATGGGTCATTGGCCAC-----TAATACCATCAGGCTGTGATCTCATTTGG	3619
Db	3658	TCGATGATGCGCCCCCATTTGGCATTAACCGGTGCAATPAAGACTCACTGGCTTTGG	3717
Oy	3620	GAATACCTAGCCCATGTCCGTAAATATGACATTTATCTTTGTCATGTTTAGACCCCGTACA	3679
Db	3718	GGTTTCCCGCCCAAGGACG---GTACGACCTGGGTTCATCACTTGGAACTAAATACA	3774
Oy	3660	GGAAACATCATACCAACAGTGGGAGGATCAGCTATCCACACAGCATGTACGTGA	3739
Db	3775	GAACACACACTTTTCAGCAGTGCAGAGACCATGGGCGACTTAACAACTTTTCGGGTT	3834
Oy	3740	AGGCTGTCCACACACTGTGAACCTGGCGAATGTGTGCTATAGGGATTTGGCTTGG	3799
Db	3835	CGGCCCTGAATTGGCTTTAACCCAGAGGACCCCTCGTGTGAATCTATGTGTAACGCG	3894
Oy	3800	ATCCGGCAACCGAATAATCATCACTGGGTGGACGCTCATTTTAGTTTACCCGTGCT	3859
Db	3895	ACCGCAACAGTAGAGACGTAGTCAACCGCTTTCAGAAAGTTTGTACAGGGTGTCTGAG	3954
Oy	3860	GTCAAGCTTAAGAACACTGCCGAATAATCTAGGTCTCTTCGTGTTCTTCCGCAAGACA	3919
Db	3955	CGAGACCGAATTTGTCTCAAGCAATACAGAAATGTACCTGATTTTCCGACACTAGCA	4014
Oy	3920	ACGGCAACCAACATATGACAGGACAGACTCGGTGTAGTGCCTTGACAACTATAT-CAA	3978
Db	4015	ACAGC-CGTACACGGCAATTCACCCCGCACATGTGAATTCGTGATTTGCTCGTGTAT	4073
Oy	3979	GGGTCAACACAGTACGAGCAGGAGAGCTCAGCGTACAGAGTATCAGGTGACATT	4038
Db	4074	GAGGGTACAAAGATGAGTTGGAGCGCGCGCTCATACCGCACAAAGGGAATAATTT	4133
Oy	4039	AGCAAGACGCTGACCAAGCTATGTGTAATCTGTATAGCAAAAGTCACACAGTTCC	4098
Db	4134	GCTGCTGTCAAGAGGAAGCAGTTGTGCATCCGACCAATCCGCTGGGTAGCCAGGCGAA	4193
Oy	4099	GGAGTGTCCGCTGCACTGTATCCGAATAATGCGCGCTCTTTGATAGACAGCCATAGCT	4158
Db	4194	GGAAGTCCGCTGCATCTATTAACAGTTGGCCGACAGTTTACCAGTTACCAACGAG	4253
Oy	4159	GTCCGGAACGGTAGACT--TGTAAGACGAAACCGCTCATCATATGCTGTAGAACCC	4215
Db	4254	ACAGGCAACCGAAGATGATGTGTGCTAGGAAGAAAGATGATCCACGCGTGGCGCT	4313
Oy	4216	AATTTTTTAAGATGCCGGAACCGGAGGCGCACTTAAAGTCCGACGCTGCTTACATGAC	4275
Db	4314	GATTTCCCGGAACACCCGAAAGCAGAAAGCTTGAATTTGTCAAAACCGCTTACATGCA	4373
Oy	4276	ATAGGTCATCTGTCACACGCTGAGCGGATTCAAATAATATCAGTACCGCTACTGTCAAC	4335
Db	4374	GTGGCAGACTTAATTAATGAACATTAACATCAAGTCTGTGCCATTTCCACTGTATCTACA	4433
Oy	4336	GGCATCTATTTCTGTGGCAAAAGATCGATGTGCAATCATTTGCATCTACTGCT	4395
Db	4434	GGCAATTTACGACGCGGAAAGAACCGGCTTGAATATCACTTAACCTGTTGACAAACGCG	4493
Oy	4396	TTCCGACTACGATGCCGATGTCAACCATATATTGCTTGGATTAACAATGGGAGACAGG	4455
Db	4494	CTAGACAGAACTGACGCGAGTGAACCATCTATTTCCTGTGATTAAGATGGAGAAAGA	4553
Oy	4456	ATATTCGAGGCGATTCACCGCAAGAAAGGTGCAAAATTCG--GGATGATGACAGGCA	4512
Db	4554	ATCGACGGGCACTTCCAACTTAAGAGTCTGTAAACAGCTGAAAGATTAAGATATGAG	4613
Oy	4513	GTAGACATTTGACTGTACGGGTCCACCAACAGCTCTTGGACGCGACGACAGTTAC	4572
Db	4614	ATCGACGATGATTTAGTATGATTCATTCACAGAGTTGCTTGAAGGGAAGAAAGGATTC	4673
Oy	4573	TCCGTCATGAGGCAAGTTGATTTCAATCTGGAAGTACAGATTCATCAGACGCGC	4632

Db 4674 AGTACTACAAAAGAAATGTATCTGACTTCGAGGACCAAAATTCATCAAGCAGCA 4733
Qy 4633 AAGACATTTGCCGAATCATGATGAGCCCAAAATCTGAGGCTATAGACATTT 4692
Db 4734 AAACATGCGGGAGATTAAGTCTCTGTTCCCTAATGACGAAAGTAATGAACACTG 4793
Qy 4693 TGTCTTACATCTCGGGGAGAGATATGTCAGCATCCGCTCCAAATGCCAGTAGAGAG 4752
Db 4794 TGTCTTACATATTTGGGTGAGACATGGAACAAATCCGCAAAAGTCCCGGTGACAT 4853
Qy 4753 TCAGAGCGTGTGCTCCACCTCAACACTTACCTGCTGTATTAATAGCTATGACGCT 4812
Db 4854 AACCGGTCTGTACCGCCGCAAAACGTTGCCGTCTTGCATGTATGACATACGCA 4913
Qy 4813 GAGCGCTATACAGATGTCGCTCTGCGAAGAAAGAGTTCGCGTATGCTATCATTC 4872
Db 4914 GAAAGGTCCACACACTTAAGAGCAATACGTCAAGAAAGTACATGATGCTCCACCC 4973
Qy 4873 CTGTTGCCAAGTACAGATCAAGCGCTGAGAGAGTACAGTCAAGCAACAGTCTG 4932
Db 4974 CCCCTCTTAAGCACAAAATTAAGATGTTCAAGAGGTTCAAGTCAAGAAAGTGTCTG 5033
Qy 4933 TTTTCAGCGCTGTACCAACCGGCTGTACACCCCGAAGATTCGCGGAAT 4982
Db 5034 TTATATCGGACACTCCCGCATTCCTCCCGCCGTAAAGTACATAGAAAT 5083

RESULT 11

US-09-507-362-103
Sequence 103, Application US/09507362
Publication No. US20030096397A1
GENERAL INFORMATION:

APPLICANT: Dubensky Jr., Thomas W.

Polo, John W.
Belli, Barbara A.
Schlesinger, Sondra
Dryga, Sergey A.
Frolov, Ilya

TITLE OF INVENTION: RECOMBINANT ALPHAVIRUS-BASED VECTORS
WITH REDUCED INHIBITION OF CELLULAR MACRO-MOLECULAR
SYNTHESIS

NUMBER OF SEQUENCES: 125

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed Intellectual Property Law Group PLLC
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA

ZIP: 98104-7092
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/507,362
FILING DATE: 18-Feb-2000

CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McMasters, David D.
REGISTRATION NUMBER: 33,963

REFERENCE/DOCKET NUMBER: 930049.457D6 /1196.011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 11740 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 103:
US-09-507-362-103

Query Match 26.0%; Score 1300.4; DB 11; Length 11740;
Best Local Similarity 56.0%; Pred. No. 0;
Matches 2819; Conservative 0; Mismatches 2111; Indels 80; Gaps 16;

Qy 23 ATATGAAAGAAATTCACGTTGACTTGAATGCTGACAGCCGCTATGTCAAGTGTACAGC 82
Db 64 AGAAGCCAGTAGTAAAGCTAGACGTAGACCCCGAGAGCGTTGTGTGCACTGCAAA 123
Qy 83 GGAGCTTCCACATTTTGAATGAGTGAACCAAGGCGATCTGACAAATGACATGCGCATG 142
Db 124 AAACCTTCCGCAATTTGAGGTAGTACACAGCGGTCACTCCAAATGACCATCTAATG 183
Qy 143 CCAGAGCTTTTGCATGTGGCAACAAGCTCATTTGAGAGGAAGTGCACCGGAGCAAG 202
Db 184 CCAGAGCATTTTGCATGTGGCAACAAGCTCATTTGAGAGGAAGTGCACCGGAGCAAG 243
Qy 203 TTATCTTGAACATTTGAAGTGCAGCGCTGACAGATGACATTTCCAAATCACCGCTATCT 262
Db 244 CGATCTTGAACATTTGAAGTGCAGCGCTGACAGATGACATTTCCAAATCACCGCTATCT 303
Qy 263 GTATCTGCTCTATGATAGAGGCTGAAAGCCGGAACAAGTCAACAGGTTATGCAAAAGC 322
Db 304 GTGTCTGCTCCCATGCTGATGTCAGAAAGACCGGACCGCATGATGAAATACCGCATGAAC 363
Qy 323 TTAAAGAA-----GTACATTTACGCAAGAAATAGCTCTTAAGCGGCGAGACCTGC 376
Db 364 TGGCGAAAGAAAGCTGCAAGATTTACAAACAAAGATTTGATAGAGATTTAGAGATTTCC 423
Qy 377 TGAAGTCAATGTCAACACAGACGACAGACTCCATCTGTGTATGACACAGACCA 436
Db 424 GGACCGTACTGATAGCGCGGAGTGTGAACACCATGCTGTCTTTCACAGCATGTGA 483
Qy 437 CGTGTAGTACTTTGGAAGTGTGACGATATCCAAAGTGTGACGACTCCATCACCGA 496
Db 484 CTGTCAACATGCGCGCAATATTCCTCATGACGAGAGTGTA---TATCAACCTCCCG 540
Qy 497 CATCAATCTCAACACAGCGCTTAAGAGATTTAGAGAAATTTAGATGAGCTTTGCA 556
Db 541 GAATATCTATCATCAGGCTATGAAAGCGTCCGAGCCCTGTATGATTTGGCTTCGACA 600
Qy 557 CGACCCCTTTTATGTACAAAACATGCGAGATTTCTTACCTTACTTACAAACGAAGTGG 616
Db 601 CCACCCAGTTCATGTTCTCGGCTATGCGAGTTGTACCTCGGCTACAAACGAAGTGG 660
Qy 617 CTGACGAGAGATTTGGAAGCAGTAACTTGGCTTCGTAATCTCAAGTCTTGAAGGA 676
Db 661 CCGACGAGAAAGTCTTGAAGCGGCTAACTGGAATTTGACAGCAAAAGCTGAAGTGAAG 720
Qy 677 GCAGGCTTGAAGAACTCAATCCTTGAAGAAAGAGGCTCCAACTTAATAAGATCA 736
Db 721 GTAGAGCAGAAATTTGTCAATATGAGAAAGAGATTTAGAGCCCGGCTCGGGGTTT 780
Qy 737 TATCTCGTGTGTTCAACAATCTACACAGAGATAGTACTTGTAGTGGATC 796
Db 781 ATTCTTCGTAGATGACACTTTATCCAAACACAGAGCTTGTGACAGCTGGCATC 840
Qy 797 TTCCAAAGTGTTCACCTTGAAGAAAGTCTAACTTCAAGTGAATGTGGACCAATG 856
Db 841 TTCCATGCTGTTCACCTTGAAGAAAGTCTAACTTCAAGTGAATGTGGACCAATG 900
Qy 857 TCGAGTGAAGGATGATGATCAAAAGATTAAGATGACAGCCGAGATTAAGTGAAG 916
Db 901 TGAGTTCGAAAGCTAGCTAGTGAAGAAATACACATCAAGTCCCGGATCAAGGAGAAA 960
Qy 917 TTGAGAACTTGGGCTCCACATGATGCGAGGTTCTTGTAGTTGCAAAAGTCAAGATA 976
Db 961 CCGTGGGATAGCGGTTACACACATAGCAGAGGCTTCTTGTATGCAAAAGTACTGACA 1020
Qy 977 CGCTGCGGCGGAGAGGGTTTCTTTGCTGTGTGATGATGATACAGCAACTTTCG 1036
Db 1021 CAGTAAAGAGAAAGCGGTATGTTCTCTGTGACAGTACATCCCGCAGCATATATGCG 1080

1037 ATCAGATGACAGGATTTGGCAACTGACGTTAGTGTGATGACGACCAAAAACTATTGG 1096
1081 ATTCAGATGATGATGATATATGCGACGATATATCACTGACGATGACAAAACTTTGG 1140
1097 TTGGGCTCAACCAAGATTTGCTCAATGTGTAGACGCAAGAAATATCAACAATGC 1156
1141 TTGGGCTCAACCAAGATTTGCTCAATGTGTAGACGCAAGAAATATCAACAATGC 1200
1157 AGAATCTATCTATTAACGATGATGCTGCGCCAGGCGTTTTCAGATGGGCGCGTAATGTG 1216
1201 AAAATTAACCTTCTGCGCATATAGCAACAAGGTTCAAGAAATGGCTAAGAGGCAAG 1260
1217 CCAGCTTGACGACGAGAAAGAACTAGAGGGTGGGAGCGGACCTTACTATGGGCTGT 1276
1261 ATGATCTTGATTAACGAGAAATGCTGGGTACTAGAGAACGCAAGCTTATGCTATGCT 1320
1277 GCTGGGCTTTCAGAACCCAGAAATCAATCTCAAGAAAGCTGTATCGCAACAA 1336
1321 TGTGGGCTTTCGACATTAAGAAAGTACATTCGTTTATGCGCCGACCTGAAACGACCT 1380
1337 TTAAAGAAAGTACCTGCGCTCTTGTGACTATTTGATTCACAGCTTATCAAGACGAGG 1396
1381 GCGTAAAGTCCAGCCTCTTTTATGCGCTTTCATGCTGCTGATGAGCACTCTT 1440
1397 TCGATATGGGCTTCCGCGTAGGCTCAAGCTGCTGCTTGAACCA--ACTGTCAACCG 1453
1441 TGCCCATGTGCTGAGGCGAATAATGAACTGGCATTTGCAACCAAGAGAGAGAAAC 1500
1454 CACCGGCTATTAACAATGCGGATGTGAGACATCTGCGGCTTACAGCAAGAAAGCTGAG 1513
1501 TCGTGCAGGCTCTCGAGGAATTAAGTCAATGAGGCAAGGCTGCTTTGAGGATGCTCAG 1560
1514 AATGTGCTGAGGAGAGATCAAGAAAGCCGCGCACCTTCTCTCCGTAATGAAA 1573
1561 AGAAGCGACAGCGGAGAACTCCGAGAACACTTTCACATTAATGTGAGCAACAAAGCA 1620
1574 AAGAGACCTGAGAGCA-----GAAATGACCTCATTTATGCAAGAGGAGAGAGCAG 1624
1621 TCGAGGAGCGCGAGAAAGTGTCTGCGAAAGTGAAGGGGCTCCAGCGCGAATCGAGCAG 1680
1625 GTAGCGTGAAGACCAAGAGACATCAAGGTTGACAAAGTTAACCGGCGAAAGAGAA 1684
1681 CATTAAGTTGAAACCCCGCGGTCACTAAGATTAATCACTCAAGCAAAATGACCGTATGA 1740
1685 TTGGGCTTACGCTATCTTCAACCCAGCGGATTAATGATGAAATCTGGCGTGA 1744
1741 TCGGACAGTATATGTTGTCTCGCAAACTGTGTGAAGAAATGCGAACAG 1800
1745 TCACACCATTTGGCGGAAACAAGTACTGATGATGACTCAAAAGGTAGGCGAGGATACA 1804
1801 CGGACCGGCTAGCAGATCAAGTTAAGTCAATACACACTCCGGAAGATCAAGAAAGTACG 1860
1805 AAGTGAAGCATTAACAAGTAAAGTCAATGTAACGAAAGGACCGCGGCTCTGTTCAAG 1864
1861 CGGTGCAACATTAAGAGCTAAAGTATGATGTCAGAGAGAGGCGGTATGAGCAG 1920
1865 ACTTCAAGGATTAAGTGAAGGCTTAAGTCTGTTTCAAGAGAGGAAATTTGTAACA 1924
1921 AATTTCTAGACATGAGTGAAGGCGCACTTAAGTGAACAAGAAAGAAATTTGTGAAC 1980
1925 GATACCTGACACATGCAATCAACGAGAGGCTTAACACTGACGAAAGATATATA 1984
1981 GCAAACTATACCATTTGCTCATGATGCCCCCGCAAGAAATCAAGAAAGAGACATACA 2040
1985 AGACTGTAAGACTCAGAGACAGACTCAGAAATACGTTTCAATATTGACGCAAGAAAGT 2044
2041 AGGTTACAAAGCAGAGCTTGCAAGAAACAGATACGTTTGAAGTGAAGAAAGCGTT 2100
2045 GTGTTAAGCAGAGAGAGCTGCTTGTGCTTAACCGGTGATCTGTGATTCACCAT 2104
2101 GCGTTTAAAGAAAGAGAGCTCAGGTGTGTTCTCTGAGAGAACTGACCAACCTCTCT 2160
2105 TTACAGAGTTTGTGTAAGAGATCTCAAGACAGACAGAGACACTCAAAAGTCCAA 2164

2161 ATCATGAGGTAGCTGTGAGAGGACTGAAGACCGGACCTGCGGTCCTGTAAGAGTCAAAA 2220
2165 CCATCGAGCTATGAGAGTCCAGGTTCAAGTAAATCTGGAATCATCAAAAAGCGCTGTGA 2224
2221 CAATGAGAGTATGAGCAACCGGGGTCCGGAAGTACCTATTAATCAAGTCAACTGTCA 2280
2225 CTAAAGAAAGTCTGTTGTGAGTGCAGAAAGAAATCGCGCAAAATCATCAGGATG 2284
2281 CGGACAGAGTCTTTGTTTACCGAGGAAAGAAAGAAATTTGTCCGAAATTTGAGCCGACG 2340
2285 TAAGAGATGAGACGATATGATGTTGCTGTAAGACTGTGATGCAATGAGTCTTAATG 2344
2341 TGCTTAAGTACGAGGGTATGCAATATACGTGAAAGCAATGATGATGCTATCTCAAG 2400
2345 GGGTTAAGCAACCCGTTAACAATCTGTATCATGATGAGGATTTGCTGCTGACAGGGA 2404
2401 GATGCCACAAAGCCGTAGAAAGTGTATGATGACCAAGGTTGCGGTGACAGCAGAG 2460
2405 CGCTGCGGACATGATTCGATGTCAAAAC--TAAGAAAGTGTATTGTGCGGGAGC 2461
2461 CACTACTGCTTGTGATTTGCTATGTGACAGGCCCCGCAAGAGTATGATATGCGGAGACC 2520
2462 CAATAATGCGGCTTCTTTAATGATATGCTGAAAGTACATTTTAAAC----- 2512
2521 CCATGCAATGCGGATTTCTTCAATGATGCAATTAAGGATCAATTTCAATCACTGAAA 2580
2513 ATGACATATGACATGAAAGTATACATTAAGAAAGCTCTGTAGAGTGCACACATGTGA 2572
2581 AAGACATATGACACCAAGACATTTACAAAGTATATCTCCGCGCTTGCACACACAGTTA 2640
2573 CCGCATGCTGTCCAGCTCTTCTACAGCAAGGAAATGAAGAGGTTAACCATGTCG 2632
2641 CAGCTATTGTATGACATCTGATTAAGATGAAAGATGAAACCAAGAACCTGCAAG 2700
2633 ATAAATCATATAGATTAACCAAGAGACCAAAAGCCGCAAAAGATGATCTGATTTGA 2692
2701 AGAAATGTAATGATATTAACAGGGGCCCAAAAGCCGAAGCCAGGAGATATCATCTGA 2760
2693 CCGTTTCAAGAGATGGGTGAAGAACACTACATGATTAACAAAATCAAGAAATCATGA 2752
2761 CATGTTTCGCGGGGTGTTAAGCAATTCAAATGACATATCCGGAATGAAGTATGA 2820
2753 CTGCGGTGATGCAAGGACTTACCGGGAAGGCGTTATGCTCAGGTAACAAGTCA 2812
2821 CAGCGCGGCTTACAAAGGCTTACCAAGAAAGAGATGTATGCGGTCGCGCAAAAGTCA 2880
2813 ACGAATCCACTCTACTCGCAGACTTGTAGACAGTGAACGTATTAATTACAGCAG 2872
2881 ATGAAACCCACTGTACGGATCAATCAAGAGATGAAACGTGTGCTCACCCGACATG 2940
2873 AAAAGCATTTGCTGGAAGACGCTAGCTGTATCCCTGATTAAGAACATTTACAGCTA 2932
2941 AGGACAGGCTAGTGTGAAGAACTTTCAGGCGGACCACTGATTAAGACAGCCCATACA 3000
2933 AATATCCGCGGATTTACGCGCTTCAATGGAAGACTGCGAGCGCAACAGCGCATTA 2992
3001 TACTTAAGGAATCTTTCAGGCTACTATAGAGACTGGAAGTGAACACAGGAAATTA 3060
2993 TGGCAGCGTCTTGTAAAGCCGACAGACGCTGATGTTCCAGAAATAGTGAACGCTT 3052
3061 TTGCTGCAATTAACAGCCCACTCCCGTGCCTCAATCGTTACAGCTCAAGACCAAGCTTT 3120
3053 GCTGGCGGAAGGCTTTAGAGCCAGTCTTGCGCACCGGCAACATTTGTGCTGACAGACAG 3112
3121 GCTGGCGGAAGGATTTGAACCGATCTAGCCACCGGCGGTATCTTAACCGGTTGCC 3180
3113 AGTG---GAGAGTTGACCCCATTAAGCATGACAGAGGATCTACCTGAATGAGCAC 3169
3181 AGTGAGCGAACTGTTTCCACAGTTTGGAGATGACAAACATTTGGCCATTTAGCGCT 3240
3170 TGAATCTTTTTCACAGAGTCTTTTGAATGACCTGACAGTGGGATATTTTCCGCTC 3229

Db 3241 TAGAGTAATTTGCAATTAAGTTTTCGATGAGCTTGAACGCGACTGTTTCTTAAC 3300
Qy 3230 CTACCGTCGCACTTACTTACAG-----GGATCAGCACTGGGATA 3268
Db 3301 AAGAGATCCCACTTAAGTACCATCCCGGATTCAGCGAGCCGGTAGCTCAATTGGAGCA 3360
Qy 3269 ACTCGCAGGAGAAACATGTAATGGCTTAATAGAGAGTACCAAGAGTTGTCAAGC 3328
Db 3361 ACAGCCAGGAGAACCGCAGTATGGGTACGATCAGCCCATTCGCCGCAACTCTCCGTA 3420
Qy 3329 GATATCCGTGATCACAAGAGGTTGACACAGCGAGGTAGCTGATTAATAGGAATATA 3388
Db 3421 GATTTCCGGTGTTCAGCTAGC---TGGGAAGGGCACACACTTGTATTCAGACGGGGA 3477
Qy 3389 CATCAAGGACTACTCTCCCAATTAATGTGTTTCATTAATGCGCGTTGCCCACT 3448
Db 3478 GAACCAAGATTAATCTCTGACAGCATTAACCTGATCCCGGTGAACGCAATCTTCCACG 3537
Qy 3449 CGTTATCGTTGACCAAAAGGACAGGGTACAATGATCAGCGGATTCATCTAAGA 3508
Db 3538 CTTATGTCCTCCAGTACAGAGAGCAACCGGCCGCTCAAAAATTTCTGAACCACT 3597
Qy 3509 TGAAGGCAATCTGTGTGTGATC---GGGATCTATCAGCATTCAGGAGAAAG 3565
Db 3598 TCMAACACCACTCAGTACTTGTGTATCAGAGAAATAATGAAGCTCCCGTAAGAGA 3657
Qy 3566 TAGAGTCATGGGTCATTTGCCAC-----TAATACATCAGGTGTATCTCGATTGG 3619
Db 3658 TCGAATGATCGCCCGATTTGGATAGCCGGTGCAGATAGAACTACAACTGGCTTTCG 3717
Qy 3620 GAATACCTAGCCATGCGTAAATATGACATTAATCTTGTCAATGTTAGACCCCTACA 3679
Db 3718 GGTTCGCCGCGAGGAG---GTAAGACTGTGTATCAACATTTGAATTAATACA 3774
Qy 3680 GGAACATATCTACCAACAGTGCAGAGATCAGCTATCCACCAAGCATCTAACGTGA 3739
Db 3775 GAAACACCACTTTCAGCAGTGCAGAACCATATGCGCGCACTTAAACCTTTGGCGTT 3834
Qy 3740 AGGCTGTCCACCACTGAACACTGGCGGAACTGTGTGCTATAGGGTATGGGCTTTCG 3799
Db 3835 CGGCTTGAATTCCTTAACCAAGAGGACCTGTGTGAATCTATATGCTACGCG 3894
Qy 3800 ATCGGCAACCGAGATATCATCTGCGGTGACGCTCATTTAGGTTAACCGCTGT 3859
Db 3895 ACCGAAAGTGAAGAGCTAGTACCGCTCTTCCAGAAATTTGTACAGGGTGTGCGAG 3954
Qy 3860 GTACGCTTAAAGCACTGCCGAAATTAAGAGTTCTTGTGTCTTTCGCAAGACA 3919
Db 3955 CGAGACAGATTTGTCTCAAGCAATTAAGAAATGTAATTTCCGACAACTAGACA 4014
Qy 3920 ACGGCAACCAACATATGACCAAGACAGCTGGGTATGTTGTAACAATCTAT-CA 3978
Db 4015 ACAGC-CGTACACGGCAATTCACCCCGCACATCTGAATTTGGTATTTTCCTGTAT 4073
Qy 3979 GGGTCAACCGAGTACGAGCAGAGAGAGCTCAGCGTACAGAGTATCAGAGTACAT 4038
Db 4074 GAGGGTACAAGAGTGAAGTTGAGCCGCCGCTATACCCGACCAAAAGGAGAAATTT 4133
Qy 4039 AGCAAGAGCGTACCAAGCTATCTGTAATGCTGTATATGAAAAGTCAACAGTTCC 4098
Db 4134 GCTGCTGTCAAGAGAAAGTGTCTCAACGACCAATCCGCTGGTATAGACAGCGCA 4193
Qy 4099 GGAGTGGCGGTGACGTGTAACGAAATGGCCGGCTGTTTGTATAGACAGCAATAGT 4158
Db 4194 GGAGTGGCGGTGACGTGTAATGAAAGTGGCCGACCAAGTTTACGATTCAGCCGAG 4253
Qy 4159 GTCCGAGCGGTAGACT---TGTGAAGACGAACCGCTCATCATCATCTGTAGACCC 4215
Db 4254 ACAGCAGCGGAGATGACTGTGTGCTAGAGAAAGATGATCCACCGCGTCCGCCCT 4313
Qy 4216 AATTTTCTTAAGATCCCGGAACCGAGGGCGACTTAAAGCTGCAGCTGCTACATGAGC 4275
Db 4314 GATTTCCGGAAGCACCCAGAAAGCAAGGCTTGAATTTGTAACAAAGCCTACATGCA 4373

Qy 4276 ATAGGTCATGTCACAGCTGAGCGGATTAACAAATATCAGTACCGCTACTGTCAAC 4335
Db 4374 GTGGCAGACTTAATGAATGAACATCAATCAAGTGTGCGCATTCACATCTATCTACA 4433
Qy 4336 GGCATCTATTCTGTGGCAAGATGAGTATGCAATCATTCATCACTGTTTCACTGCT 4395
Db 4434 GGCATTTACGAGCGCGGAAAGACCGCTTGAAGTATCACTTAATGCTGTGACACCGG 4493
Qy 4396 TTGCACTAGGATGCCGATGTCACTATATTTCTTGATTAACAAATGGAGACGAG 4455
Db 4494 CTAGCAAACTGACGCGGACGTAACTATATTCCTGATTAAGAGTGAAGAGAAAGA 4553
Qy 4456 ATATCGAGGCGCATTCACCGCAAGAAAGCGCGAATTTCT---GGATGATGACAGCA 4512
Db 4554 ATCGAGCGGCGACTCAATTAAGAGTCTGTATACAGAGCTGAAAGATGAATATGAG 4613
Qy 4513 GTAGCATTTGATTTGTGAGGATCCACCAACAGCTTTTGGCAGGACAGCAAGTTAC 4572
Db 4614 ATCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4673
Qy 4573 TCCGCTAATGAGGCAAGTTGATTCATCTGAGAGTACAGATTCATCAGACCGC 4632
Db 4674 AGTACTACAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4733
Qy 4633 AAGGACATTTGCCGAATTCATGCAATGAGGCGCAACAAATCTGAGGCTAATGAGAGATT 4692
Db 4734 AAGGACATTTGCCGAATTTAGGCTGTTCTTCTTAATGACAGAAAGTAAAGAACTG 4793
Qy 4693 TGTCTTATCATCTGCGGAGAGTATGTCTCAGCATCCGCTCCAAATGCGCAGTAGAG 4752
Db 4794 TGTGCTTATATTTGGTGTGAGACATGGAAGCAATCCGCAAAAGTGTCCGCTGACCAT 4853
Qy 4753 TCAGAGGCTGTCTTCACTTCAACACTTCATGCTGTGTATTAATGCTATGAGGCT 4812
Db 4854 AACCGGTGTAGCGCCCGCAAAAGCTGTGCTTGTGATGATGATGATGATGATGATGAT 4913
Qy 4813 GAGCGGTATCAGGTTGCGCTTCCGAAAGAAAGAGTTCGCGTATGCTATCATTC 4872
Db 4914 GAAAGGTCACAGCTTAAGAAAGCAATACGTCAAGAAAGTAAAGTATGCTCTCCAC 4973
Qy 4873 CTGTTCCGAAGTACAGATCAGAGCGTGCAGAGATCACTACAGTGCAGCAACAGTCTG 4932
Db 4974 CCCCTCTTAAGCAAAATTAAGATGTTTGAAGATGTTTGAAGTGCAGAAAGTATGCTG 5033
Qy 4933 TTTTCAAGCGTGTACACCGGCTGTACACCCAGAAAGTACGCGGAAT 4982
Db 5034 TTTAATCCGACACTCCCGCATTCGTTCCCGCTGATGATGATGATGATGATGATGAT 5083

RESULT 12
US-09-507-362-101
; Sequence 101, Application US/09507362
; Publication No. US20030096397A1
GENERAL INFORMATION:
APPLICANT: Dubensky Jr., Thomas W.
; Polo, John W.
; Belli, Barbara A.
; Schlesinger, Sandra
; Dryga, Sergey A.
; Prolov, Ilya
TITLE OF INVENTION: RECOMBINANT ALPHAVIRUS-BASED VECTORS
WITH REDUCED INHIBITION OF CELLULAR MACRO-MOLECULAR
SYNTHESIS
NUMBER OF SEQUENCES: 125
CORRESPONDENCE ADDRESS:
ADDRESS: Seed Intellectual Property Law Group PLLC
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/507,362
FILING DATE: 18-Feb-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 930049.457d6 /1196.011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 101:
SEQUENCE CHARACTERISTICS:
LENGTH: 8000 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 101:
US-09-507-362-101

Query Match 25.9% Score 1295.6; DB 11; Length 8000;
Best Local Similarity 56.0%; Pred. No. 0;
Matches 2816; Conservative 0; Mismatches 2134; Indels 80; Gaps 16;

QY 23 ATATGAAAGAAATTCACGTTGACTTGAATGCTGACAGCCCGTATGTCAGTTCGTTACAGC 82
DB 64 AGAAGCCAGTAGTAACGTAGACGTAGAGACCCCAAGATCCGTTGTGCTGACAGTGAANA 123
QY 83 GGAGCTTCCCAATTTAGATTCGAAGCAGAGCAGTCTACTGACAAATACCATGCCAATG 142
DB 124 AAACCTTCCCAATTTAGATTCGAAGCAGAGCAGTCTACTGACAAATACCATGCCAATG 183
QY 143 CCAGAGCGTTTTCGATGTGGCAACAAAGCTCATTTGAGAGCGAAGTCCAGCGGAGCAAG 202
DB 184 CCAAGGCAATTTCCGATCTGGCCAGTAACTAATGAGCTGAGGTTCTTACACAGGA 243
QY 203 TTATCTTGGACATTTGGAAGTGCAGCCCGTCAAGCATGACATTCATTCATCCAGCTAAT 262
DB 244 CGATCTTGGACATTTGGAAGTGCAGCCCGTCAAGCATGACATTCATTCATCCAGCTAAT 303
QY 263 GTATCTGCTTATGATTAAGCGCTGAAGACCCGAGACGACTACAAAGGATATGCAAGAAC 322
DB 304 GTGCTGCTTATGATTAAGCGCTGAAGACCCGAGACGACTACAAAGGATATGCAAGAAC 363
QY 323 TTAAGAAAA-----GTGACATTTACCGAAGAAATAGCTCTTAAGGCGGAGACTGCG 376
DB 364 TGGCGGAAAAAGCGTGCAAGATTACAAACAAAGATTGATGAGAAATTAAGGATCTCC 423
QY 377 TGAAGTATGATCAACACAGACGAGAGACTCCATCTCTGTATGACACAGACGCA 436
DB 424 GGACCGTACTGATATGCGCGGATGCTGAACACCATCTCTGCTTTCACAAAGATGTTA 483
QY 437 CGTGTAGTACTTTGGAAGTGTAGCATATACCAAGATGTGACCAAGTCCATGACCGGA 496
DB 484 CCGTGAACATCGCGTCCGAATATTTCCGTATGACAGAGCGTGA---TATCAACGCTCCG 540
QY 497 CATCAATATACACAGGCGCTTAAGAGATTAGACATTTACTGATAGGCTTTGACA 556
DB 541 GAACTATATATATAGGCTATGAAAGCGTGCGGACCTGTACTGATGATGCTTTCGACA 600
QY 557 CGACCCCTTTTATGATCAAAAAACATGAGAGTTCTTACCTACTTACAAACAGAGCTGG 616
DB 601 CAACCAATATATGATTTCTGCGTATGAGCAGGTTGCTACCTGCGTAAACACCAACTGG 660
QY 617 CTGACGAGAGATTTAGGACAGCATTAATTTGCGTCTGCTGATCACTCAATTTTCAAGGA 676
DB 661 CCGACGAGAAAGTCTTGAAGGCGCTTAACATCGGACTTTGACGACCAAAAGCTGAGTGAAG 720
QY 677 GCAGGCTTGAAGAACTCTCAATCTTAGAAGAAAGGCTCCAACTACTAATAGATCA 736

DB 721 GTAGGACAGAAAAATGTGCATATATAGGAAAGAGGTTGAAGCCCGGGTCCGGGTTT 780
QY 737 TATTTCTGGTTGGTTCAACATATCAACAGAAATATGATACATGTTATCGTATCGGCATC 796
DB 781 ATTTCTCCGTAGATTCACATTTATTCACAAACACAGAGCAGCTTGCAGAGCTGGATC 840
QY 797 TTCCAAAGGTTTCCACTTAAAGAAAGTCTAACTTACAGAGTATGTGGGCAATTTG 856
DB 841 TTCCATGCGTGTTCACCTTAAAGAAAGAGTGTACATTTGCGCTGTGATACATG 900
QY 857 TCAAGCTGAAAGGATTCATCAAAAAGATTAACGATCAACCCAGACATACGTTAAG 916
DB 901 TGAGTTCGAAGGCTACTGATGTAAGAAATTCACATCAAGTCCCGGATCAAGGAGAAA 960
QY 917 TTGAAACTTGGGCTCCACATGATGCGAGGGTTTCTTGAATGTCAGAAAGTCAAGATA 976
DB 961 CCGTGGGATACCGCGTTTACACAAATGACAGGGCTTCTGCTAATGCAAAAGTTACTGACA 1020
QY 977 CGCTGGCGGCGAGAGGGTTTCTTTGCTGTGTATGATATGATACAGCCACACTTTGG 1036
DB 1021 CAGTAAAGGAAACGGGATATCGTTCCCTGTGTGACATGATCCGGCCACATATGCG 1080
QY 1037 ATCAGATGACAGGATTTCTGCAACTGACGTTAGTGTGATGACGCAAAAACTATTGG 1096
DB 1081 ATCAGATGACGATGATATATGAGCCAGCATATATCACTGACGATGACAAAACTTCTGG 1140
QY 1097 TTGGGCTCAACCAAGGATTTGCTCATGTGTAGAGACGCAAAAGAAATCTACCAATGC 1156
DB 1141 TTGGGCTCAACCAAGGATTTGCTCATGTGTAGAGACGCAAAAGAAATCTACCAATGC 1200
QY 1157 AGAATCTATATTAACGATGATGCGCGGAGGCTTTCCAGGTGGGCGCGTGAACATGCTG 1216
DB 1201 AAATTTACCTTGTCCATCTATGACACAGAGGTTCAAGAAATGGCTTAAGAGCGCAAG 1280
QY 1217 CCGACTTGAACGACGAGAAAGAACTAGAGGCTGCGGAGCGCATCTTATATGAGGCTGCT 1276
DB 1261 ATGATCTTGAATTAAGAAAGAAATGCTGGTACTAAGAAACGCAACTTACGATGCTGCT 1320
QY 1277 GCTGGCTTCAAGACCCAGAAATCAATCATCTACAAAGCTGTGTACGCAACMA 1336
DB 1321 TGTCGGGCTTTCGCACTAAGAAAGTACATTCGTTTATGCGCCACTGGAACCGAGACT 1380
QY 1337 TTAAGAAAGTACCTGCGGCTTTGACTCATTTGTGATTTCCAGCGCTTACAGCCAGGCG 1396
DB 1381 GCGTAAAGTCCAGCGCTTTTATGCGCTTTTCCATGTGTGCTATGACAGACTCTT 1440
QY 1397 TCGATATGAGGCTTCCGCGTATAGGCTCAAGCTGCTTGAACCA---ACTGTCAACCCG 1453
DB 1441 TGCCCATGTGCTGAGGAGAAATTTGAACCTGGCATTTGCAACCAAGAAAGGAGAAAAC 1500
QY 1454 CACCGGCTATTAATGATGCGCGATGTGAGCATCTGCGGTTCACAGCAAGAACTGAAG 1513
DB 1501 TGCTGCAAGTCTTCGAGAGAAATTAATGATGAGGCGCAAGGCTCTTTTGAAGATCTCAG 1560
QY 1514 AAGTGGCTGACGAGGAGAAAGATCAGAAAGCCGCGCACTTGTCTCCCTGAATTAAGAA 1573
DB 1561 AGAAGGCGAGGCGGAGAAAGCTCCGAGAGGACATTTCACTTATGTGCAAGCAAAAGCA 1620
QY 1574 AAGAGCGGTGAGAGCA-----GAAGTGAACCTCATTTATGCAAGAGCAGAGAGAG 1624
DB 1621 TCGAGGACGCGCGAGAAAGTTGTCTGGAAGTGTGAGAGGGCTCAGAGCGGACATTCGAGAG 1680
QY 1625 GTACGCTGAGACACACAGAGAGACATCAGGTGTGACAGTTACCCAGGCGAAGAGAA 1684
DB 1681 CATTAATTAAGAAACCCCGCGGTGACGTAAGGATTAATCTCAAGCAAAATACCGTATGA 1740
QY 1685 TTGGGTCTTAAGCATATCTTCAACCCAGGCGGATTAATGAGAAAACTGGCGTGA 1744
DB 1741 TCGACAGTATATGTGTTGTCTGCGCAAACTTGTACTGAAAGATGCCAAATCTGCCAGAG 1800
QY 1745 TCCACCAATTTGGCGAAACAGTACTGTTAATGACTCAAAAGTTAGGCGCAGAGATCA 1804

Db 1801 CGCACCCGCTAGACATCAGGTTAAGATCATTAACACTCCGGAAATCAGGAAGTACG 1860
Qy 1805 AAGTCAGACCATATACCGAGTAAAGTCAATTGTACAGAAAGGACGGCGGTCCCTGTTCAAG 1864
Db 1861 CGGTGAACCATACGACGCTTAAGTATGATCCACGACGAGAGTCCCTGATCCATGCGCAG 1920
Qy 1865 ACTTCAGGCAATTGAGTGAAGACGGCTACGATGTTTTCAACGAGAGGAGTTTGTAAACA 1924
Db 1921 AATTCTGACACTGAGTGAAGCGCCAGCTTAGTATACAGAAAGAGGCTGTGAACC 1980
Qy 1925 GATACCTGCACCACTCCGCAATCAACGAGAGCGCTAAACCTGACGAAAGTACTATA 1984
Db 1961 GCAAACTATACCACTTCCCATGATGCGCCCGCAAGATACAAAGAGGAGCGTACA 2040
Qy 1985 AGACTGTAAAGACTCAGACACAGACTCAGAAATACGTCTTGATTTGACGACGAAAGT 2044
Db 2041 AGGTTACAAAGGACAGCTTGACGAAACAGAGTACGTGTTGACGTGACAAAGACGTT 2100
Qy 2045 GTGTTAAGCGAAGACCGAAGTCCCTTGTCCTAACCGGTGATCTGTAGATCCACAT 2104
Db 2101 GCGTTAAGAGGAAAGACCTCAGGTCGTGCTCTCCGGAGAACTGACCAACCTCCCT 2160
Qy 2105 TTCAGAGATTGGCGTACAGAGTCTCAGACACGACGACAGCACTCAAAAGTCCAA 2164
Db 2161 ATCATGACCTAGCTCTGAGAGGACTGAAAGACCCGACTGCGGTCCCTACAGAGTCAAA 2220
Qy 2165 CCATCGAGTCTATGAGTGCAGGTTCCAGGTAATCTGAAATCATCAAAAGCGCTGTA 2224
Db 2221 CATATGAGTATGACACACCGGGGTGCGGCAATCTATATCAAGTCACTGTCA 2280
Qy 2225 CTAAAGAAAGATCTGTTGTGAGTGCAGAAAGAAACCTGCGCAGAAATCATCAGGATG 2284
Db 2281 CGGACGAGATCTTTTACAGCGGAAAGAAAGAAATTTGCGCAATTTAGGCGCAG 2340
Qy 2285 TAAAGAGATGACGATGATGATGTTGCTGCTAGAGCTGTGATTCAGTCTTAAATG 2344
Db 2341 TGTAAAGCTAGAGGGGTATGAGATTACGTGAAACAGTAGATTCGGTTATGCTCAAG 2400
Qy 2345 GGGTTAAGCACCCTGTAACACTGTATCATTTGATGAGGCAATTCCTGCCATGAGGA 2404
Db 2401 GATGCCAAGAGCCGTGAGAGTGTGATGAGGAGCGTTCCGTCACACGAGAG 2460
Qy 2405 CGCTGCTGCACTGATTCATCTCTCAAAAC--TAAAGAAAGTGTATTTGTCGGGAGC 2461
Db 2461 CACTACTTGCTGTATGCTATCTGTCAGGCCCGGCAAGAGTATATCTATGCGGAGAC 2520
Qy 2462 CAAACCAATGCGGCTTTTAAATGATGCTGTAAGATCATTTTAAAC----- 2512
Db 2521 CATGCAATGCGGATCTTCAACATGATGCACTAAAGTACATTTCAATCACCTGAAA 2580
Qy 2513 ATGACATATGACATGATGATACATTAAGCATCTTAGAGGTGCAACAGACTGTA 2572
Db 2561 AAGACATATGACCAAGACATTTCTAACAGTATATCTCCGGGTGTCACACGCGATTA 2640
Qy 2573 CCGGCATGCTCTCCAGCTCTTCTACGACAAAGCATGAAAGAGGTTAACCATGTGCTG 2632
Db 2641 CAGCATTTGTATCGACACTGATTTACATGAGAAATGAAACACAGAACCCGTGCAAGA 2700
Qy 2653 ATAAATCATCATATAGATACACAGGAGCACAAAGCCGCAAAAGATGATCTGATTTCA 2692
Db 2701 AGAACATTAATCATATTTACAGGGGCGCAAAAGCCAAAGCGAGGATATCATCTGTA 2760
Qy 2693 CCTGTTTGAAGAGATGGGTGAAGACGCTACAGATTGATACAAATAATACGAAATCATGA 2752
Db 2761 CATGTTTCCGGGGGTGAAGCAATTGCAATGACATATCCCGAATGAAAGTAAGA 2820
Qy 2753 CTGCGGCTGATCGAGAGACTTACGCGAAAGCGTTTATGCTCTCAGATCAAAAGTCA 2812
Db 2821 CAGCGCGGCTCTCAAGAGGCTTACAGAAAGAGGTGATGCTCCGGGCAAAAGTCA 2880
Qy 2813 ACAGATATCACTTACTGACAGACTTTCAGACAGTGAACGTGTTACTTACACGACAG 2872
Db 2881 ATGAAACCACTGATGCGATCATCATGAGAGCATGTGAACGTGTTGCTCACCCGCACTG 2940

Qy 2873 AAAAAAGCATTTGCTGGAAGACGCTAGCTGGTGAATCCCTGATTAAGACACTTACAGCTA 2932
Db 2941 AGGACAGGCTATGTGTGAAAACTTTCAGGGCGACACCTGATTAAGACAGCTACTAACA 3000
Qy 2953 AATATCCCGGGGATTTTCCAGGCTTCATTTGACAGCATGCGGACGCGCAACAGACGCTTA 2992
Db 3001 TACTTAAGGAAACTTTCAGGCTACTATAGAGAGCTGGGAAAGCTGAAACAAAGGAATTA 3060
Qy 2993 TGGCAGCGTCTTGTATAGCGCGACAGACGTGATGTGTTCCAGAAATTAAGTGAACGCT 3052
Db 3061 TTGCTGCAATTAAGACGCCCACTCCCGTGCATTCGTTCACTGCAAGACCAACGTTT 3120
Qy 3053 GCTGGGCGAAGGCTTTAGAGCCAGTCTGGGCGACGCGCAACATTTGCTGACAGACAGC 3112
Db 3121 GCTGGGCGAAGGATTTGAAACCGATCTAGCCACGCGCGGTATGCTACTTACCGTTGCC 3180
Qy 3113 AGTGG--GAGACGTTGCAACCATTTCAAGCATGACAGAGCGTACTCACCTGAATGGAC 3169
Db 3181 AGTGGAGGAATGTTCCACAGTTGGGATGACAAACCAATTCGCGCATTTACGCT 3240
Qy 3170 TGAATCTTTTGGACCAAGTCTTTTGGATGAGACTGGAACAGTGGGTTATTTCCGCTC 3229
Db 3241 TAGAGTAAATTTGATTAAGTTTTCGCTAGTGAACAGGCGACTGTTTCTTAAC 3300
Qy 3230 CTACCGTGCACCTTACTTACAG-----GGATCAGACATGGGATA 3268
Db 3301 AGAGATCTCCATTAACATGATCCCGGATTTACGACGAGCGGTGACTATTTGGACA 3360
Qy 3269 ACTGCGCAGGAGAAACATGATGGGCTTAATAGAGATGACAAAGAGTTGTCACGCG 3328
Db 3361 ACAGCCAGGAACCGCAAGTATGGGTACATCAACGATTCGCGCGAATCTCCGTA 3420
Qy 3329 GATATCGTGATCAAAAGCGGTTGACACAGGAGGTGCTGATTAAGAAATATA 3388
Db 3421 GATTTCCGTTTCCAGTAC--TGGGAAGGCGACAACTGATTTGACAGCGGGA 3477
Qy 3389 CCATCAAGAGACTACCTCCCAATTAATGATTTCACTAAATGCGCGTTCGCCACT 3448
Db 3478 GAAACAGATTAATCTCTGACACATTAACATGATGATCCGTAACGCAATCTCTCACG 3537
Qy 3449 GGTGATGTTGACACCAAGAGACAGGGTCAATGATCAGAGCGATTCCTATTAAGA 3508
Db 3538 CTTATGTCGCGAGTACAGAGGAAGCAACCGCGCGGTGAAAAATTTTGAACAGT 3597
Qy 3509 TGAAGGCAATCTGTGTGTGATC--GCCGATCTATCAGACTTCAGGAAAGAAAG 3565
Db 3598 TCAAAACACACTCAGTACTTGTGTATCAGAGAAATAAGTCTCCGTAAGAGA 3657
Qy 3566 TAGAGTCATGGGTCCATGCGCCAC-----TAATACATGAGGTGATCTGCAATTTGG 3619
Db 3658 TCGAATGATGCCCCGATGCGATGAGCGGTGACATTAAGACATCAACCTGGCTTTCG 3717
Qy 3620 GAATACCTAGGCATGCGGTAAATATGACATTAATCTTTGTCATATGTTAGAACCCCGTACA 3679
Db 3718 GGTTCGCGCGCAGGACG--GTACAGCTGTGTGTTTATCAATTTGGAATTAATACA 3774
Qy 3680 GGAACCATCACTACCAAGTGCAGAGATCAGCTATTCACCAACAGATGCTTAACGTGTA 3739
Db 3775 GAAACACCACTTTACAGAGTGCAGAGACATGCGGCGACCTTAAAAACCTTTTCGCGTT 3834
Qy 3740 AGGCTGTCACCACTGAAACCTGCGGAAATATGATGTTGCTATAGGATGAGGCTTGTG 3799
Db 3835 CGGCTCTGAATTTGCTTATCTAGAGAGCATCTGCTGTGTAAGTCTTATGCTACGCG 3894
Qy 3800 ATTCGCGCAACGAGAAATATCATCACTGCGGTGAGACGCTCATTTAGGTTTACCGGTCT 3859
Db 3895 ACCGCAACAGTGAAGAGTATCAACCGCTTTCGCAAGAAATTTGTCAAGGTGTCTGAG 3954
Qy 3860 GTACAGCTTAAGAAACATGCGGAAATATCTAGGTTCTTGTGTTCTTTCGGAAGACA 3919
Db 3955 CGAGACAGATTTGTCTCAAGCAATACAGAAATGATCTGATTTTCGACCACTAGACA 4014

OY	727	TATTCGCGGTGGTTCACAAATCTACACAGAGATAGATCACTGTGACGTAGCTGGCATC	796
Db	1673	ATTCTTCGGTAGGAATCGACACTTTATTCGAAACACAGAGCCAGCTTGACGACGTGGCATTC	1733
OY	797	TTCCAAACGCTGTTCACCTTGAAGAAGAAAGTCTTAACCTTCACAGGTAGATGTGGACCATTG	856
Db	1723	TTCCATCGGTGTTCACCTTGAATGGAAAGCACTGTAACCTTGGCGCTGTGTATACAGTGG	1792
OY	857	TCAGCTGTGAAGGGTAGCGTCATCAAAAAGATAGATCAGCCACAGACTATACGGTAAAG	916
Db	1793	TGAGTTGGCAAGGCTACGTAGTGAAGAAAATACCATCAAGTCCGGGATACAGGGAGAAA	1855
OY	917	TTGAGAACTTGGCGCTTCACAAATGCATCCGAGGGTTTCTTGAGTTGCAAAGTCACAGATA	976
Db	1853	CCGTGGAGTACCGCGGTTCACAAATAGGAGGGCTTCTTGCTATGCAAAAGTTACTGCACA	1912
OY	977	CGCTCGCGGCGAGAGGGGTTCTTTTGGCTGTGTAGATGTATACAGCCACACTTTCG	1033
Db	1913	CAGTAAAAGGAGAACGGGTATGTTCTTCCGTGTGACGTACATCCCGGCACACATATCG	1972
OY	1037	ATCAGATGACAGGGATTTCTGGCACTGACGTTAGTGTGGATACCCACAAAACTAATTGG	1096
Db	1973	ATCAGATGACGTGTATTAATGGCCAGGATATATCACTGACGATGCACAAAATCTTCG	2033
OY	1097	TTGGGCTCAACCAAGGATTTGCTGCATATGTAGACGCAAGAAATCTAACACATATGC	1156
Db	2033	TTGGGCTCAACCAAGGAAATGTCAATTAACGTAGAGACTAACAGGAACACAAACCATATGC	2092
OY	1157	AGAACTATCTATTTACAGTGGCGGCCAGGGCGTTTCCAGTGGGGCGCGTAAACATCGT	1216
Db	2093	AAATATACCTTCTGCCATCATAGACAGAGGTTCAACAAATGGGCTAAGGCGCCAAAG	2153
OY	1217	CCGACTTGGACGACGAGAAAGAACTAGGGGTGCGGAGCGCACTTTACTATGCGCTCT	1276
Db	2153	ATGATCTTGATTAACGAGAAAATGCTGGGATCTAGAGAAACGAAAGTTACGTATGGCTCT	2212
OY	1277	GCTGGGCTTTCAAGACCCAGAAAATCAATCCATCTACAGAGCCTTGTAACGCAACAA	1336
Db	2213	TGTGGGCTTTGCGCACTAAGAAAGTACATTCGTTTATTCGCCCACTGGAAACGACAGACT	2272
OY	1337	TTAAGAAAGTACCTGCGCTTTGACTATTTGTATTCACAGCCTTACAGCCACGGGG	1396
Db	2273	GGCTAAAAGTCCCAAGCTCTTTAGCGCTTTCATGTCGTCCGTATGACGACTCTT	2332
OY	1397	TGATATGGGCTTCCGCGTAGGCTCAAGCTGCTCTGAAACCA---ACTGTCAAAACCG	1453
Db	2333	TGCCATGTCTCGTAGGCGAAGAAATTGAAACTGTGCATTGCACCAAGAAAGAGAAAAC	2392
OY	1454	CACCGGCTATTAACAATGGCCGATGTGAGCATCTGCGGCTTACAGCAAGAGCTGAAG	1513
Db	2393	TGCTGCAAGGTCTCGAGGAATTAATCATGAGGCGCAAGGCTGTGTTGAGATGTCAAG	2452
OY	1514	AAGTGGCTGCAAGCGGAAGAGATCAGAGAAAGCCTGCAACCTTGTCTCCTGAATGAAG	1573
Db	2453	AGGAAGCCAGAGCGGAGAAAGCTCCGAGAGCACTTTCACACTTATGTGCAAGCAAGGCCA	2512
OY	1574	AAGAGCCGTAGAGCA-----GAAGTAGACCTCATTAATGCAAGAGGACGAGAG	1624
Db	2513	TCGAGGACCGCGCAGAAAGTGTCTGCGAAGTGGAGGGCTCCAGCGGCAATCGAGCAG	2572
OY	1625	GTAAGCTGAGACACACAGAGGACACATCAAGGTGACAAGTTACCCAGCGAAGAGAGA	1684
Db	2573	CATTATGTTGAACCCCGCGGCTACGTAAAGATTAATCTCTACAGCAAAATGACGCTAAGA	2632
OY	1685	TTGGGTCTTAAGCTATACTTTCACCCCAGGGGTATTGAATATGAAAAATCTGGCTGT	1744
Db	2633	TCGACAGTATATCGTGTCTCGCCAAACTGTGTCTAAGAAATGCCAAATCTCGACAG	2692
OY	1745	TCACCCATTGGCGGAGCAAGTACTGTATATGACTCAAAAGGTGGGACGAGGAGATACA	1804
Db	2693	CGACACCCGCTAGCATATAGTTAAGATCATTAACAACCTCCGGAAGATCAAGAAAGTACG	2752
OY	1805	AAGTCAGACCATACACGTAAGGTCAATGTATACAGAAAGGACGCGGCTCCCTGTTCAAG	1864

Db	2753	CGGTGAAACCATGACCCCTAAAGTACTGATGCCAGACGGAGGTGCCGTACCATGGCCAG	2812
Qy	1865	ACTTCACAGGCATTTGATGAGAGCCGTACGATCGTTTTCAACGAGAGGGAGTTCTTAACA	1924
Db	2813	AATTCTTAGCACTGATGAGAGCCGACGTTAGTGTACAAAGAAAGAGAGTTGTGAACC	2872
Qy	1925	GATACCTGACCACTCCCAATCAACGGAGAGCGCTTAACACTGACGAGAGTACTATA	1984
Db	2873	GCAAACTTATCCACTTCCCATGTGATGCCGCCCAAGAAATACAGAAAGAGACGATACA	2932
Qy	1985	AGACTGTAAAGACTCAGGACACAGCTGAGATAGCTTCTTGATATTGACCCAGAAAGT	2044
Db	2933	AGGTTACAAAGCAGAGCTTGACAAAMACAGATACGTTTGACGTGGACAAAGCGTT	2992
Qy	2045	GTGTTAACCGAAGAACGACAGGTCCCTTGTCCTCAACCGGTGATCTGGTAGATCCACAT	2104
Db	2993	GGCTTTAAAGAGAAAGAACCTCAGGTCTGTCTCTCTGGAGAACTGACCACTCTCCCT	3052
Qy	2105	TTCAAGATTTGCGTACGAGAGTCTCAAGACAGACGACGACCTCACAAAGTCCCA	2164
Db	3053	ATCATGAGCTAGCTCTGGAGGGACTGAAGACCCGACCTGCGGTCCCGTACAAAGGTCGAA	3112
Qy	2165	CCATCGGAGTCTATGGAAGTCCAGGTTCAAGTAATCTGAAATCATCAAAAGCCCTGGA	2224
Db	3113	CAATAGAGTATATGGCAACCCGGGGTCCGGCAAGTCAGCTATTATCAAGTCACTGTCA	3172
Qy	2225	CTAAGAAAGATCTGGTTGTGAGTGCAGAAAGAAACCTGCGCAAGAAATCATCAGGAGTG	2284
Db	3173	CGGACCGAGATCTTGTTHACAGCGGAAAGAAAGAAATTTGTCCGAAATTTGAGGCCAGAC	3232
Qy	2285	TAAGAGAGATGACGATATGAGATGTCTGCTAGAGCTGTGATTCAGTGTCTTAAATG	2344
Db	3233	TGCTTAAGACTAGGGGATATGCAATTAAGTGAAGACAGTAGATTCGGTTATGCTCAACG	3292
Qy	2345	GGGTTAAGACACCCGTTAACTCTGTATCATTTGATGAGCAATTTGCTGCCATGACAGGA	2404
Db	3293	GATGCCAAMAAACCGTATGAAAGTGTGTGATGCGTGAAGAGGTTGCGTGCACACAGAG	3352
Qy	2405	CGCTGCTGGCACTGATTGGCCATGCTGTCAAAC---TAAGAAAGGTATTTGGCGGGACC	2461
Db	3353	CACTACTTGCTTATTTGCTATGCTGACGCCCCGCAAGAAAGTATGTACTATGCGGAGACC	3412
Qy	2462	CAAAACAATGCGGCTCTTTAAACATGATGTGCTGAAAGTACATTTTAAAC-----	2512
Db	3413	CAATGCAATGCGGATTTCTTAACATGATGCACTAAAGTATCATTTCATCACTCCGTAAA	3472
Qy	2513	ATGACATATGCACTGAAGTGTACCATTAAMAGCATCTCTAGAGAGGTGCACACAGCTGAA	2572
Db	3473	AAGCATATATGACCAAGACATTTCAACAAGTATATCTCCCGGCTGTGCACACAGCAATTA	3532
Qy	2573	CCGCACTGTGTCCACCGCTCTTTACAGACAAAGCAATGAAAGACGTTAAACCCATGTGCTG	2632
Db	3533	CAGCTATTGTATGCACACTGCATTACGATGAGAAAGATGAAACCAACGAAACCCGTGCAAG	3592
Qy	2633	ATTAATATCATATGATATACCAACAGGACCAACAAAGCCGCAACAAAGATGATCTGATTTCAA	2692
Db	3593	AGAACTTTGAATATGATATTTACAGGGGCGCACAAAGCCGAAAGCCGAGGGGATATCATCTTGA	3652
Qy	2693	CTGTGTTACAGAGATGGGTGAAACAGTACAGATTTGACTACAAAATCACAGAAATCATGA	2752
Db	3653	CATGTTTCCGGGGGTGGTTAAGCAATTTGCAAAATGCACTATCCGGGACATGAAGTATATGA	3712
Qy	2753	CTGGCGGTGATTCGCAAGAGATTACGGGAAAGGCCGTTTATGCTGTCAAGTACAAAGTCA	2812
Db	3713	CAGCCGGCGGCTTCAAGAGGCTTAACGAAAGAGAGTATGCGCTCCGGCAAAAGATCA	3772
Qy	2813	ACGAGAAATCCACTTACTTCGAGACTCTTGAGACGTTGAACGTTTACTTACACGACAG	2872
Db	3773	ATGAAAAACCACTATAGCCATTCACATCAAGACATGTGAACGTTGCTTCAACCCGACTG	3832
Qy	2873	AAAAAGCATTTGCTGAAAGCGCTAGCTGTGATATCCGTGATTAAGACACTTACAGCTA	2932

Db 3833 AGGACAGCTAGTGTGAAAACTTTCAGGCGGACCCATGATTTAAGACGCCACTAACA 3892
Qy 2933 AATATCCGGGGATTTTCAGCGCTTCATTGACGACTGGACGGGAAACAGAGCCATT 2992
Db 3893 TACTTAAAGGAATTTTCAGGCTACTATAGAGACTGGAGACTGAAACAGAGGAAATTA 3952
Qy 2993 TGGCAGCGCTTCTGATTAAGCCGACAGCTGATGTGTTCAGAAATTAAGGTGAACGCT 3052
Db 3953 TTGTGTCAATTAACAGGCCCACTCCCGTGCATTCGGTTCAGTGCAGAACCAAGCTTT 4012
Qy 3053 GCTGGGGGAAGCTTTAGAGCGAAGCTTTGGCCACAGGCCACATTTGTCTGACGACAGC 3112
Db 4013 GCTGGGGGAAGCAATTGGAACCGTACTAGCACAGCGCGATGTGTCTTAACCGGTGGC 4072
Qy 3113 AGTGG---GAGACGTTCACCCATTCAAGCATGACAGGCGTACTCACTGAATGCGAC 3169
Db 4073 AGTGGACGAACTGTTCACAGTTTGGGATGACAAACACATTCGGCACTTTACGCT 4132
Qy 3170 TGAACCTCTTTTGACACAGGTTCTTTGAGTGAAGCTGACAGTGGTATTTTCGCTC 3229
Db 4133 TAGACGTAATTTGATTAAGTTTTCGGCATGACTTGACAGGCGACTGTTTCTAAC 4192
Qy 3230 CTACCGTGCATTAATTACG-----GGATCAGCATGGGATA 3268
Db 4193 AGAGCATCCCACTACCTACCATCCCGCATTCAGCGAGGCGGTAGCTATTGGACA 4252
Qy 3269 ACTCGCAGGGAAGAATGATTTGGGCTTAATAGAGGTAGCAAAAGATTTGCAGGC 3328
Db 4253 ACAGCCGGAACCCGCAAGTATGGTACGATCAGCGCATTTGCGCGAACTCTCCGTA 4312
Qy 3329 GATATCCGTGCATCAAAAACGGTTGACACAGCGAGGTAGCTGATATAAGAAATATA 3388
Db 4313 GATTTCCGGTGTTCACGTAGC---TGGAAAGGCAACAACCTGATTTGACAGCGGGA 4369
Qy 3389 CCATCAAGGACTACTCTTCAACAATTAATGTGTTCATTAAATCCGCGTTCCTCACT 3448
Db 4370 GAACCAAGTATCTCTGACACAGCATTAACCTGTCCCGGTGACCGCAATCTTCTCAG 4429
Qy 3449 CGTTGATCGTTGACCAAAAAGGACAGGGTACACTGATCAGCGGATTCCTATCTAAGA 3508
Db 4430 CTTAGTCCCGAGTACAAAGAGACAACTCCGCGCTCAAAAATTTCTTGAACAGT 4489
Qy 3509 TGAAGGCAATCTGTGTGTGTATC---GGCGATCTATCAGCATTCAGAGGAAGAAG 3565
Db 4490 TCAAAACCACTCACTGATCTGTGTATCAGAGAAAAAATTGAAGCTCCCGCTAAGGAA 4549
Qy 3566 TAGAGTCAATGGGTCTATGCCCCA-----TAATACATCAGGTGTGATTCGATTTGG 3619
Db 4550 TCGATGATCGCCCGCATTTGGCATAGCCGGTGCAGATTAAGAACTAACAACCTGCTTTCG 4609
Qy 3620 GAATACCTAGCGATGTGGTAAATATGACATTAATCTTTGCAATGTTAGACCCCGTACA 3679
Db 4610 GGTTCGCGCGGACGACG---GTACAGCTGGTGTTCATTAACATTTGAACTAAATACA 4666
Qy 3680 GGAACCTCACTACCAACAGTGCAGAGTCAAGCTATCCACACAGCATGCTTAACGTGA 3739
Db 4667 GAAACCACTTTCAAGAGTGCAGAACACATGGGGGACCTTAAAAACCTTTTCGGGT 4726
Qy 3740 AGGCTGTCAACCACTGAAACCTGGCGGAACATGTGTGATTAAGGTATGGGCTTGGC 3799
Db 4727 CGGCGCTGAATTTGTTAACTCAAGAGGACCTCTGTGTGAAATCTCATAGGCTACCGC 4786
Qy 3800 ATGGCGCAACCGAATATATCATCTGCGGTGACAGCTCATTTAGTTTACCGTGTCT 3859
Db 4787 ACCGCAACAGTGAAGAGTATGATCAGCTCTTTCAGAAAGTTTGTCAAGGTTGTCTGAG 4846
Qy 3860 GTACGCTTAAGAACACTGCGAAATATCTGAGTTCTTTGCTGTCTTTCGGCAAGACA 3919
Db 4847 CGAGACCAAGATTGTCTCAAGCAATACAGAAATGTACTGATTTTCGAACTATGACA 4906
Qy 3920 ACCGCAACCAACATGACAGGACAGACTCGGTGTATGTGCTTGAACAATCTAT-CAA 3978
Db 4907 ACAGC-GGTACACGCAATTCACCCGACCATCTGAATGCGTATTCGTCGTGTAT 4965

Qy 3979 GGGTCAACAGGTACGAGGAGAGAGCTCCAGCGTACAGAGTATCAGAGGTGACATT 4038
Db 4966 GAGGTTCAAAGATGAGATTTGAGAGCGCGCCGTCTATCCGCACAAABAGGAAATATT 5025
Qy 4039 AGCAAGCGCTGACCAAGCTATCTTAATGTCTTAATAGCAAAAGGTCAACAGTTCC 4098
Db 5026 GCTACTGTCAAGAGAAAGATTTGTCAAGCGACCAATCCGCTGGGTAGACAGCGAA 5085
Qy 4099 GGAATGCGGTGCATCTGACCGAAAAATGGCGGCTCTTTTGAATACAGCCAAATGCT 4158
Db 5086 GGAATGCGGTGCATCTGATTAACCTTTGGCCACCAAGTTTACCGATTACGCCAGAG 5145
Qy 4159 GTCGGAGCGCTAGACT---TGTAAAGCAAGAACCGCTCATATACATGCTGTAGAGACC 4215
Db 5146 ACAGGACCGCAAGATAGACTGTGTGTCTTGAAGAAAGATGATCCAGCGGTGCGCT 5205
Qy 4216 AATTTTCTAAGATGCCGAACCGAGGCGACCTTAAGCTCGCAGCTGCTACATGAGC 4275
Db 5206 GATTTCCGGAAGCACCGAAGACGAAGCTTGAATTTGCTTACAAAAGCCTACATGCA 5265
Qy 4276 ATAGCTTCATCTGCAACGCTGAGCGGATTAACAAAATATCAGTACCGCTACTCAAC 4335
Db 5266 GTGCGACACTTAATTAATGAACATACATCAAGCTGTCCCATTCACCTGCTATTA 5325
Qy 4336 GGCATCTATTCTGTGGCAAAAGATCGAGTATGCAATTCGATCACTGTTCACTGCT 4395
Db 5326 GGCATTTACGCAACCGAABAAAGACCGCTTGAAGTATCACTTAATCTGTGACAAACCG 5385
Qy 4396 TTGCACACTACGAGTGCATGTCAACATATTTGCTTGAATAAACATGAGAGACGAG 4455
Db 5386 CTAGACAGACTGACGCGGAGCTAACCATCTATTTGCTGTGATTAAGAGTGAAGAGAA 5445
Qy 4456 ATATTCAGGCTCATTCACCCGCAAGAAAGCTGAAATTTCT---GATGATGACAAACCA 4512
Db 5446 ATGACGCGGCTCTCAACTTAAGAGTCTGTAACAGAGCTGAAGATGAAGATATGAG 5505
Qy 4513 GTAGACATTAATGCTGTGCTGAGGTCACACCAACAGCTTTTGGCAGGACACAGGTTAC 4572
Db 5506 ATGACGATATGATTAATGATTTCAATCCAGACAGTTGCTTGAAGGAABAAAGGATTC 5565
Qy 4573 TCCGTCAATGAGGCAAGTTGATTCATACCTGGAAGTACAGATTCATCAGACCGCC 4632
Db 5566 AGTACTCAAAAGAAATTTGATTTGCTTTCGAAGGACACCAATTCATCAAGCAGCA 5625
Qy 4633 AAGGCAATTCGCAAAATTCATGCAATGTGGCCCAACAAATCTGAGGTTATGACAAATT 4682
Db 5626 AAAGACATGGCGGAGATTAAGTTCCTGTTCCTAATGACAGGAAGATTAATGAACAATG 5685
Qy 4693 TGTCTGATCTCCGCGGAGAGATATGTCAGCATCCGCTCCAAATCCAGTACAGAG 4752
Db 5686 TGTGCTTCAATTTGGTGTAGACCATGAGACATCCGCAAAAGTCCCGGTGACCAT 5745
Qy 4753 TCAAGGCGTCTGCTCCACCTCAACACTTCATGCTGTGTAATTAAGCTATGACGGCT 4812
Db 5746 AACCGGTGCTAGCCGCGCCCAAAAGCTTTCGATGTATGTCATGACGCA 5805
Qy 4813 GAGCGCTATACAGTGTGCGCTCTGCGAAGAAAGACAGTTCCCGGTATGCTCATTTTC 4872
Db 5806 GAAAGGTTCCACAGACTTAAGACAAATACGTCAAAAGAAATTCACAGTATCTCTCCACC 5865
Qy 4873 CTGTTCGGAAGTACAGGATCACAGCGGTGACAGAGTACAGTGCAGCAACAGTCTG 4932
Db 5866 CCCCTTCTTAAGCACAAAATTAAGATTTCAAGAGTTTCAAGTGCAGAAAGTATGTCCTG 5925
Qy 4933 TTTTCAGCGCTGTACCAACCGGCTGTACACCCAGGAAGTACCGGAAAT 4982
Db 5926 TTTAATCCGACACTCCCGCATTCGTTCCCGCGTAAAGTACATTAAGT 5975

RESULT 14
US-09-275-883-1
; Sequence 1, Application US/09275883

Publication No. US20030053988A1
 GENERAL INFORMATION:
 APPLICANT: Renner, Wolfgang A.
 APPLICANT: Nieba, Lars
 APPLICANT: Boersma, Marco
 TITLE OF INVENTION: Inducible Alphaviral Gene Expression System
 FILE REFERENCE: 1700.0020001
 CURRENT APPLICATION NUMBER: US/09/275,883
 CURRENT FILING DATE: 1999-03-25
 EARLIER APPLICATION NUMBER: US 60/079,562
 EARLIER FILING DATE: 1998-03-27
 NUMBER OF SEQ ID NOS: 9
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 1
 LENGTH: 11282
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: cDNA
 US-09-275-883-1

Query Match 25 %; Score 1295.6; DB 11; Length 11282;
 Best Local Similarity 56.0%; Pred. No. 0;
 Matches 2816; Conservative 0; Mismatches 2134; Indels 80; Gaps 16;

23 ATATGAGAAGATTACGCTTGAAGATGCTGACAGCCGATGCTCAAGTGGTACAGC 82
 956 AGAGCCAGTGTAACTGACGCTGACAGCCGATGCTGCTTCTGCTGACAGTGCATA 1015
 83 GAGCCTTCCCAATTTAGATGAGAGGAGGAGGAGTCACTGACAAATGACATGCAATG 142
 1016 AAGGCTCCCGCAATTTAGAGTGTAGAGCAGCAGTCACTCAATGATGCTGCTGCTG 1075
 143 CCAAGCCTTTTCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 202
 1076 CAGAGCAATTTTCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1135
 203 TTATCTGACATTTGAGAGTGGCGCTGACAGCATGACATTCATTCATTCATTCATTCAT 262
 1136 CGATTTGACATTTGAGAGTGGCGCTGACAGCATGACATTCATTCATTCATTCATTCAT 1195
 263 GTATCTGCTTATGATGAGAGTGGCGCTGACAGCATGACATTCATTCATTCATTCATTCAT 322
 1196 GTGCTGCTTATGATGAGAGTGGCGCTGACAGCATGACATTCATTCATTCATTCATTCAT 1255
 323 TTAGAAAA-----GTGACATTTACGACAGAGATGCTGCTGCTGCTGCTGCTGCTGCT 376
 1256 TGGCGGAAAAAGCGTGAAGATTTACAAACAGATTCATTCATTCATTCATTCATTCATTC 1315
 377 TGGAGTCAATGACACACAGAGAGTCCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 436
 1316 GAGCGCTTATGATGAGAGTGGCGCTGACAGCATGACATTCATTCATTCATTCATTCAT 1375
 437 CGTGTAGTACTTTGAGAGTGGCGCTGACAGCATGACATTCATTCATTCATTCATTCATTC 496
 1376 CCGTGAATGAGTGGCGCTGACAGCATGACATTCATTCATTCATTCATTCATTCATTCAT 1432
 497 CATCAATTTACACAGAGTGGCGCTTAAAGATGAGCAATTTACTGATAGGCTTTGACA 556
 1433 GAACATTTATCATGAGCTTAAAGAGTGGCGCTGACAGCATGACATTCATTCATTCATTC 1492
 557 CGACCCCTTTATGATGACAAAAACATGGAGGTTCTCACTCACTTACAAACAGAGTGG 616
 1493 CACCCAGTATGATGATGAGAGTGGCGCTGACAGCATGACATTCATTCATTCATTCATTC 1552
 617 CTGAGAGAGATTTGAGAGTGGCGCTTAAAGATGAGCAATTTACTGATAGGCTTTGACA 676
 1553 CCGAGAGAAAGTCTTGAAGGCGTAAATCGAGCTTTGAGAGCAAAAGCTGAGATGAG 1612
 677 GCAGCTTGAAGAACTCTCAATCTTGAAGAGAGAGGCTCCAACTTCAATTAAGATCA 736
 1613 GTAGAGCAGAAATTTGTGATTAATGAGAGAGAGAGGCTCCGAGTCCGAGGTTT 1672

737 TATTCCTCGTGTGTTCAACAATCTACAGAGATGATCACTGTTAGTAGCTGGCATC 796
 1673 ATTTCCTCGTGTGTTCAACAATCTTATCCAGAAACAGAGCCAGCTTGCAGAGCTGGCATC 1732
 797 TTCCAAAGTGTGTCTTCAATTTGAGAGAAAGTCTTCACTTCAAGTATGATGAGAGTGG 856
 1733 TTCCATCGTGTGTCTTCAATTTGAGAGAAAGTCTTCACTTCAAGTATGATGAGAGTGG 1792
 857 TCAGCTGGAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 916
 1793 TGAGTTGGAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1852
 917 TTGAGAACTTGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 976
 1853 CCGTGGAGTATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1912
 977 CGCTGCGCGGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1036
 1913 CAGTAAAGAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1972
 1037 ATCAGATGACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1096
 1973 ATCAGATGACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2032
 1097 TTGGGCTCAACCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1156
 2033 TTGGGCTCAACCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2092
 1157 AGAATCTATCTTATCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1216
 2093 AAAATTTACCTTCTGCGCATCTATGACACAGAGGTTTCAAGAAATGAGGCTTAAAGG 2152
 1217 CCGACTTGAAG 1276
 2153 ATGATCTTGAATGAG 2212
 1277 GCTGGGCTTCAAG 1336
 2213 TGTGGGCTTCAAG 2272
 1337 TTAAGAAATGATCTGCGCTCTTTGATCTGATTTGATTTCAAGGCTTTCAAGGAGAG 1396
 2273 GGTAAAGTCTGCGAG 2332
 1397 TCGATATGAGGCTTCCGCGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1453
 2333 TGCCATGCTGCTGAG 2392
 1454 CACGGGCTATTAATGAG 2452
 2393 TGCTGAGAGTCTGAG 2512
 1514 AAGTGTGAG 1573
 2453 AGAAG 2512
 1514 AAG 1624
 2513 TCGAG 2572
 1625 GTAGAGTGAAG 1684
 2573 CATTAATGTAAG 2632
 1685 TTGGGCTTCAAG 1744
 2633 TCGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2692
 1745 TCGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1804
 2693 CGAG 2752
 1805 AAGTGAAGCATTAACAG 1864

Db 2753 CGGTCGAACCATGACGAGCTTAAAGTATGATGCCAGAGAGGCGCTTACATGGCCAG 2812
Qy 1865 ACTTCCAGGCTTATGATGAGAGCCGCTACGATCGTTTTCACAGAGGAGATTCGTAACA 1924
Db 2813 AATTCCCTAGGCACTGATGAGAGCCGACGTTATGTTACAGAAAGAGATTTGTGAACC 2872
Qy 1925 GATACCTGCACCATGATGCAATCAACGAGAGGCGCTAAACACTGACGAAAGATCTATA 1984
Db 2873 GCAACCTATACCACTATTCGATGATGCGCCGCCAGAAATACAGAAAGAGAGCAATCA 2932
Qy 1985 AGACTGTAAAGACTCAGAGACACAGACTCGAATACGTTCTGATATTTGACGACGAAAGT 2044
Db 2933 AGGTTACAAAGGACAGAGCTTCAAGAAACAGATAGTGTGATGATGATCAAGAGCGTT 2992
Qy 2045 GTGTAAAGCCGAGAGACGAGAGTCCCTTGTGCTTACCGGTGATCTGTTATATCAACAT 2104
Db 2993 GCGTTAAGAAAGAAAGAGCCCTCAGGTCTGTCTCTCGGAGAACTGACCAACCCTCCCT 3052
Qy 2105 TTACAGATTGGCTACGAGAGTCTCAAGACAGCAGCAGCAGCCTCAGAAAGTCCCA 2164
Db 3053 ATCTAGAGCTAGCTCTGAGAGGATGTAAGACCAGCTCGCGTCCCTACAGAGTCCAAA 3112
Qy 2165 CCATCGAGTCTATGAGATGTCAGGTTCAAGTAAATCTGAATCATCAAAAGCGCTGTA 2224
Db 3113 CAATAGAGGTATGAGCACAACGGGGTGGGCAAGTCAAGTATATCAAGTCACTGTCA 3172
Qy 2225 CTAAGAAAGATCTGTGTTGTGATGCCAGAAAGAAAGAAAGAAAGTTCGCAAAATGAGCCG 3232
Db 3173 CGGACAGAGATCTTGTATCAACGCGAAAGAAAGAAAGTTCGCAAAATGAGCCGCA 3292
Qy 2285 TAAAGAGATGAGACGATGATGTTGCTGTAGAGCTGTGATTCAGTGTCTTCTAAATG 2344
Db 3233 TGCTAAAGACTGAGGGGATGATGAGATTAAGTGAAGACAGTATTCGTTATGCTCAAG 3292
Qy 2345 GGGTTAAGACCCCGTTAAACCTGTATCATTTGATGAGGCAATTTGCCCTGCAAGGGA 2404
Db 3293 GATGCCCAAAAGCCGTGAAGATGCTGTACGTTGAGAAAGCGTTGCGGACCAAGAG 3352
Qy 2405 CGCTGCTGCGACTATGTTGCCATCTGCAACC---TAAAGAAAGTGTATTTGGCGGAGC 2461
Db 3353 CACTACTGCTCTTATGTTGATGCTATGCTCAGGCCCCGCAAGAAAGTATGATTAAGCGAG 3412
Qy 2462 CAAACATGCGGCGCTCTTAAACATGATGCTGAAAGTACATTTTAAC----- 2512
Db 3413 CCAATGCAATGCGGATCTTCAACATGATGCACTTAAAGTACATTTCAATCAACCTGAAA 3472
Qy 2513 ATGACATATGCACTGAAGTGTACATTAAGCAATCTTAAAGAGTGCACACAGACTGTA 2572
Db 3473 AAGACATATGCAACAGACATTTCTAAGATATATCTCCGCGTGCACACAGCCAGTTA 3532
Qy 2573 CCGGCAATGCTCCACGCTCTTCAACGACAGCAAGCAATGAACGGTTAAACCATGTGCTG 2632
Db 3533 CAGTATATGATTCGACACTGCAATTCATGATGAAGATGAACACAGAACCCGTCAGAA 3592
Qy 2633 ATAAATCATCATGATATCACACAGGACCAAAAGCCGCAAAAGATATCTGATTTCTAA 2692
Db 3593 AGAACATTTGAATATGATTTTCAAGGGGCAAAAGCCGCAACCCGAGGATATCATCTGTA 3652
Qy 2693 CTTGTTTCAAGATGAGTGAAGACGCTACAGATTTGACTCAAAAATCAAGAAATCATGA 2752
Db 3653 CATGTTTCCGGGGTGGGTTAAGCAATTCGAAATTCGCGAGCATGAAGTAATGA 3712
Qy 2753 CTGCGGCTGATTCGCAAGAGACTTACGCGGAAAGCGTTTATGCTGTCAAGTTCAAAAGTCA 2812
Db 3713 CAGCGCGGCTTCAACAGGGCTTAAACAGAAAAAGAGTGTATGCGTCCGCGCAAAAGTCA 3772
Qy 2813 ACGAATATCACTGATCTGAGACTTCTGAGCAAGTGAAGTGAATTTCTTACAGCAAG 2872
Db 3773 ATGAAACCCACTGTATCCGATCAATGAGCAATGTATGCTTCTTCCCGGCACTG 3832
Qy 2873 AAAAAGCATTTGTGAGAGAGCGTAGCTGTGATTCCTGATTAAGCACTTACAGCTA 2932

Db 3833 AGGACAGGCTAGTGTGAAAACTTTCAGGGCGAGCCCATGATTAAGCACCCCACTAACA 3892
Qy 2933 AATATCCCGGGGATTTTCAAGGCTTCAATTTGAGACACTGGCAGCGGAGACAGCCCATTA 2992
Db 3893 TACTTAAAGAAATTTTCAGGCTATATAGAGACTGGGAAGCTGAAACACAGGGAAATTA 3952
Qy 2993 TGGCAGCGCTTCTGATTAAGCCGAGACAGCTGATGTGTTCCAGAAATGAAGTGAACGCT 3052
Db 3953 TTGCTGAATTAACAGCCCACTCCCGTCAATCCGTTCAAGTGCAGAACCAAGCTTT 4012
Qy 3053 GCTGGCGAAAGGCTTATGAGCCAGCTTTGGCCACAGCCCAATTTGTCTGACAGACAGC 3112
Db 4013 GCTGGCGAAAGCAATTTGAAACCGATTAACAGCCGCGGTATCTGATTAACCGGTTGCC 4072
Qy 3113 AGTGG---GAGAGCTTCCACCATTTCAAGATGACAGAGGCTACTCACTGAAATGGCAC 3169
Db 4073 AGTGAAGCAACTGTTTCCACAGTTTGGATGACAAACCACTTTCCGCCATTTACGCT 4132
Qy 3170 TGAACCTTTTTCACAGGCTCTTTGAGATGACCTGACAGTGGTTATTTTCCGCTC 3229
Db 4133 TAGACGTAATTTGATTAAGTTTTTCGCAATGACCTTGACAGGCGACTGTTTCTAAAC 4192
Qy 3230 CTACCGTGCATTTACTTAAC-----GATCAGCACTGGGATA 3268
Db 4193 AGAGCATCCCACTAACCTATCCCGCATTCAGCAGGACCGGTAGCTATTGGGACA 4252
Qy 3269 ACTGCGAGGAAAGAAATGATGATGAGCTTAATGAGAGTGAAGCAAGTGTGCAGCGC 3328
Db 4253 ACAGCCAGAAACCCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4312
Qy 3329 GATATCCGTCATCACAAAAGCGGTTGACACAGCAGGATGATGATTAAGAAATATA 3388
Db 4313 GATTTCCGGTGTTCACATAGC---TGGAAAGGCAACAACTGATTTGACAGCGGGA 4369
Qy 3389 CCATCAAGACTACTCTTCAACATTAATGTTGTTCCATTAATCCCGGTTGCCCACT 3448
Db 4370 GAACCAAGTATCTCTGACAGCATTAACCTGTCGCGTGAACCGCAATCTTCTCAG 4429
Qy 3449 CGTTGATGTTGACCAAAAGGACAGGTTCAACTGATCAACGCGATTCCTTCTAAGA 3508
Db 4430 CTTAGTCCCGAGTACAAAGAGACACACCCGCGCTCAAAAATTTTGAACAGT 4489
Qy 3509 TGAAGGCAATCTGCTGTGATGATC---GCGATCTTATCAGATTTCAAGGAAAGAAAG 3565
Db 4490 TCAACACACACTAGTACTTGTGTATCAGAGGAAATTTGAAGCTCCCGTAAAGAA 4549
Qy 3566 TAGAGTCCATGCGCTCTTGTGCCAC-----TAATACATCAGGTGATCTCGATTTGG 3619
Db 4550 TCGAATGATTCGCCCCGATGAGATGACCGGTGCAATTAAGAACTCAACCTGCTTTGG 4609
Qy 3620 GAATACCTAGCAATGTCGTTAAATATGACATTTATCTTGTCAATGTTAGGACCCCGTACA 3679
Db 4610 GGTTCGCGCGGAGCAG---GTACGACCTGTGTTCATCACTGAACTTGAATTAATCA 4666
Qy 3680 GGAACCATCACTACCAACAGTGCAGAGATCAAGCTATCAACACAGCATCTAACTGTA 3739
Db 4667 GAACACACACTTTCAACAGTGGAAAGACATGCGGACCTTAAAAACCTTTCCGCTT 4726
Qy 3740 AGGCTGTCCACACCTGAACACTGCGGAAACATGTGTCTATAGGATATGCGCTTCTG 3789
Db 4727 CCGGCTTGAATTTGTTTAACTCAGAGGCAACCTCTGCTGGAAGTCTTAATGCTAACCCG 4786
Qy 3800 ATGCGGCAACCGGAATATCATCATGCGGTGGAGCGCTATTTAGGTTTACCCTGTCT 3859
Db 4787 ACAGCAACATGAGAGCTGATGCTACGCTCTTGCAGAAAGTTGTACAGGTGTCTCAG 4846
Qy 3860 GTACCTTAAGAAACACTGCGGAAATCTGAGGTTCTTCTGTTCTTCCGCAAGACA 3919
Db 4847 CGAGACAGATTTGTCTCAAGCAATACAGAAATGATACGATTTTCCAGCACTAACA 4906
Qy 3920 ACGGCAACCAACACATGACAGAGCAAGTCCGTTGATGCTTTGACACATCTAT-CAA 3978
Db 4907 ACAGC-CGTACAGGCAATTCACCCGCACTGTAATGATGATGATGATGATGATGATGAT 4965

QY 3979 GGTCAACAGGTACAGGAGGAGAGCTCCAGCTACAGATGATCAGAGTGCAT 4038
DB 4966 GAGGGTACAGAGATGAGATGGAGCCGCCGTATACCGACCAAGAGAGATAT 5025
QY 4039 AGCAAGAGCGGTGACCAAGCTATCGTTATGCTGCTAATAGCAAGAGTCAACAGGTTCC 4098
DB 5026 GCTGACTGCAAGAGAGACAGTGTGCAACGACGCCAATCCCTGGGTGAGACAGGCGAA 5085
QY 4099 GGAGTGTGGGTGACCTGTAACGAAATAGCCGGGCTGTTGATGACAGCAATAGCT 4158
DB 5086 GGAGTGTGGGTGACCTGTAATGAGCTTGGCCGACAGTTTACGATTCAGCCACGGAG 5145
QY 4159 GTCCGAGCGGTGACT--TGTGAAGCAGAACCGCTCATATACATGCTAGGACCC 4215
DB 5146 ACAGGACCGGAGATGACTGTGCTAGAGAAAGATATCCACCGGTCGGCCT 5205
QY 4216 AATTTTCTAGATCGGAGACCGGAGGCGACTTAACTGCGAGCTGCTACATGAC 4275
DB 5206 GATTTCCGAGACACCCAGAGACAGAGCTTGAAATTGCTCAAAACCGCTACATGCA 5265
QY 4276 ATAGGTCCATGCTCAAGCTGAGCGGATTAACAAATTCAGTACCGCTACTGTCAAC 4335
DB 5266 GTGGAGACTTATGTAATGAAATACATCAACAGTCTGTCCATTCACATGCTATCTACA 5325
QY 4336 GGCATCTATTTCTGTGCGCAAGATGAGTATGATGATCAATGCTGCTGCTCACTGCT 4395
DB 5326 GGCATTTAGCGACCGGAGAAAGACCGCTTGAAGTATCACTTAACTGCTTGACACCGG 5385
QY 4386 TTGCACTACGAGTGCAGATCAACATATATGCTTGATTAACATAGGAGACGAG 4455
DB 5386 CTAGACAGAACTGACGCGGACGTAACCATCTATTGCTGATGAGAGTGAAGAAAGA 5445
QY 4456 ATATCGAGCGCATTCACCGCAAAAGACCTGCAATTTCT--GATATATGACAGCA 4512
DB 5446 ATCGAGCGGCACTCCAACTTAAAGAGTCTGTAACAGAGCTAAGAGTAAATATGAG 5505
QY 4513 GTAGCATTTGATGCTGAGGCTCCACCAACAGCTCTTTGGAGGCGACAGAGTTAC 4572
DB 5506 ATCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5565
QY 4573 TCCGCTAATGAGGCAAGTGTATTCATCTGTAAGAGTACAGATTCATGACCGCC 4632
DB 5566 AGTACTACAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5625
QY 4633 AAGCATTTGCCGAATCCATGATGATGATGATGATGATGATGATGATGATGATGATG 4692
DB 5626 AAAGCATGCGGAGATTAAGTCTGTTCCCTTAATGACAGAGAAATGATGATGATGATG 5685
QY 4693 TGCCTTGCATCTGAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 4752
DB 5686 TGTGCTTACATATTTGGGTGAGACATGAGAGAAATGATGATGATGATGATGATGATG 5745
QY 4753 TCAGAGGCGTGTCTCACTCACAACCTTCATGCTGTGATTAATTAAGCTATGACGCT 4812
DB 5746 AACCGTGTGTAGCCGCCCAAAAGTTCGCTTGTGATGATGATGATGATGATGATG 5805
QY 4813 GAGCGCTATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4872
DB 5806 GAAAGGCTCAGAGCTTAAGAGCAATACGTAAGAGTGAAGTGAAGTGAAGTGAAGTGA 5865
QY 4873 CTGTTGCGAAGTACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 4932
DB 5866 CCCCCTTCTAGACAAATTAAGATGATGATGATGATGATGATGATGATGATGATGATG 5925
QY 4933 TTTTCAAGCGTGTACACCGGCTGTACACCCAGAGAGTACGCGAAT 4982
DB 5926 TTTATCGCACACTCCGCAATTCGTTCCGCCGTAAGTATGAAAGT 5975

RESULT 15
US-10-123-101-51
; Sequence 51, Application US/10123101

/ Publication No. US20030148262A1
/ GENERAL INFORMATION:
/ APPLICANT: POLO, John
/ APPLICANT: PERRI, Silvia
/ APPLICANT: THUDION, Kent
/ TITLE OF INVENTION: CHIMERIC ALPHAVIRUS REPLICON PARTICLES
/ FILE REFERENCE: 2300-17924 / PPI1924.002
/ CURRENT APPLICATION NUMBER: US/10/123.101
/ CURRENT FILING DATE: 2002-07-23
/ PRIOR APPLICATION NUMBER: 60/295.451
/ PRIOR FILING DATE: 2001-05-31
/ NUMBER OF SEQ ID NOS: 92
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 51
/ LENGTH: 705
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: VEE nsp fragment oligo 1
US-10-123-101-51

Query Match 6.6%; Score 330.8; DB 12; Length 705;
Best Local Similarity 67.0%; Pred. No. 1e-97;
Matches 470; Conservative 0; Mismatches 232; Indels 0; Gaps 0;

QY 627 AGATTGGAGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 686
DB 3 AGAGTTAAGCGCTGTAACATGAGCTTATGACCTTATGATGAGCGGTACGTAG 62
QY 687 AAACCTCAATCTCTTGAAGAGAGAGCTCCACCTTATATATGATCATATTCCTGCT 746
DB 63 AGGATGTCATTTCTTAAAGAGATATTTGAAACATCAACATGTTCTATTTCTCT 122
QY 747 TGTTCACCAATCTACACAGAGATGATGATGATGATGATGATGATGATGATGATGATG 806
DB 123 TGGCTCGACATCTACACAGAGAGAGAGTACTGAGAGCTGACCTGCTCTCT 182
QY 807 GTTCACCTTGAAGAGAGAGTCTTACTTCAAGTATGATGATGATGATGATGATGATGATG 866
DB 183 ATTTACTTACGTGCGCAAGCAAAATTAACATGCTGCTGATGATGATGATGATGATGATG 242
QY 867 AGGATGCTCATCAAAAGATTAACATGATGATGATGATGATGATGATGATGATGATGATG 926
DB 243 CGGATGATGCTTAAAGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATG 302
QY 927 GCGTTCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 986
DB 303 TGTGCTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 362
QY 987 CGAGAGGCTTTCTTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1046
DB 363 GAGAGGATCTTTTCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 422
QY 1047 AGGATTTGCAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1106
DB 423 TGGCATTTGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 482
QY 1107 CCAGAGATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1166
DB 483 CCAGGATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 542
QY 1187 ATTACAGTGTGCGCCAGGAGGTTTTCAGATGAGGCGGTAAACATTCGTCCGATTTGA 1226
DB 543 TTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 602
QY 1227 CGACAGAGAAAGTATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1286
DB 603 AGATGAAAGGCACTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 662
QY 1287 CAAGACCCAGAAATACATCATCTTACAGAGAGCTGTGATC 1328
DB 663 TAGAAGGACCAAGATTAACATCTATTTATAGCGCCCGGATAC 704

Mon Nov 17 10:39:22 2003

us-10-023-649-1_copy_1_5000.rnpb

Page 38

Search completed: November 16, 2003, 00:28:09
Job time : 1549 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 12:07:42 ; Search time 308.Seconds
(without alignments)
7165.311 Million cell updates/sec

Title: US-10-023-649-1_COPY_1_5000
Perfect score: 5000
Sequence: 1 accctacaactacatgcgc.....ataattcagaacgcacc 5000

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA:*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/6C_COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4492	89.8	11492	3	US-08-991-840A-1
2	2118	42.4	11459	3	US-09-454-721A-3
3	2074.8	41.5	11464	3	US-08-991-840A-2
4	1564	31.3	11517	1	US-07-920-281C-1
5	1564	31.3	11517	3	US-08-466-277-1
6	1564	31.3	15538	4	US-09-554-337-1
7	1473.8	29.5	8100	4	US-09-554-337-4
8	1333.2	26.7	11663	1	US-08-446-932-1
9	1333.2	26.7	11663	1	US-08-801-263A-1
10	1333.2	26.7	11663	1	US-08-801-263A-7
11	1333.2	26.7	11663	3	US-09-102-248-1
12	1333.2	26.7	11663	3	US-09-102-248-7
13	1333.2	26.7	11663	4	US-09-367-764-1
14	1333.2	26.7	11663	4	US-09-367-764-7
15	1303.6	26.1	11703	1	US-08-801-263A-8
16	1303.6	26.1	11703	3	US-09-102-248-8
17	1303.6	26.1	11703	4	US-09-367-764-8
18	1302.2	26.0	11717	1	US-08-801-263A-4
19	1302.2	26.0	11717	4	US-09-102-248-4
20	1302.2	26.0	11717	4	US-09-367-764-4
21	1300.4	26.0	8000	4	US-09-415-784-102
22	1300.4	26.0	8000	4	US-09-415-785A-102
23	1300.4	26.0	8000	4	US-08-944-465-102
24	1300.4	26.0	8000	4	US-09-415-868-102
25	1300.4	26.0	8000	4	US-09-415-900-102
26	1300.4	26.0	11740	4	US-09-415-784-103
27	1300.4	26.0	11740	4	US-09-415-785A-103

28	1300.4	26.0	11740	4	US-08-944-465-103	Sequence 103, App
29	1300.4	26.0	11740	4	US-09-415-868-103	Sequence 103, App
30	1300.4	26.0	11740	4	US-09-415-900-103	Sequence 103, App
31	1298.8	26.0	9951	4	US-09-193-707-3	Sequence 3, Appli
32	1298.8	26.0	10524	4	US-09-193-707-4	Sequence 4, Appli
33	1298.8	26.0	11927	4	US-09-193-707-5	Sequence 5, Appli
34	1298.8	26.0	13905	3	US-08-972-218-1	Sequence 1, Appli
35	1298.8	26.0	13905	3	US-09-193-707-1	Sequence 1, Appli
36	1295.6	25.9	8000	4	US-09-415-784-101	Sequence 101, App
37	1295.6	25.9	8000	4	US-09-415-785A-101	Sequence 101, App
38	1295.6	25.9	8000	4	US-08-944-465-101	Sequence 101, App
39	1295.6	25.9	8000	4	US-09-415-868-101	Sequence 101, App
40	1295.6	25.9	8000	4	US-09-415-900-101	Sequence 101, App
41	1295.6	25.9	11282	4	US-09-733-042-1	Sequence 1, Appli
42	1294	25.9	16656	1	US-08-741-881-1	Sequence 1, Appli
43	1294	25.9	16656	1	US-08-739-158-1	Sequence 1, Appli
44	1294	25.9	16656	2	US-08-739-157-1	Sequence 1, Appli
45	1294	25.9	16656	3	US-08-404-796-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-991-840A-1
; Sequence 1, Application US/08991840A
; Patent No. 6261570
; GENERAL INFORMATION:
; APPLICANT: Michael D. Parker
; APPLICANT: Jonathan F. Smith
; APPLICANT: Bruce Crise
; APPLICANT: Mark Steve Oberste
; APPLICANT: Shannon Schumura
; TITLE OF INVENTION: Live Attenuated Virus Vaccines for Eastern Equine Encephalitis
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Charles H. Harris
; STREET: USA MRC - MRC-JA
; CITY: FORT DETRICK, FREDERICK
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.5
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,840A
; FILING DATE: December 16, 1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Provisional Application 60/047162,
; FILING DATE: May 20, 1997
; APPLICATION NUMBER: Provisional Application 60/053,652
; FILING DATE: July 24, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles H. Harris
; REGISTRATION NUMBER: 34,616
; REFERENCE/DOCKET NUMBER: 003/058/SAP RIID 96-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 619-2065
; TELEFAX: (301) 619-5034
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11492 base pairs
; TYPE: Nucleic acid
; STRANDEDNESS: Double
; TOPOLOGY: Linear
; FEATURE: OTHER INFORMATION: N at all occurrences is = unknown:
; FEATURE: OTHER INFORMATION: K at all
; FEATURE: occurrences is = G or T
US-08-991-840A-1

Query Match 89.8%; Score 4492; DB 3; Length 11492;
Best Local Similarity 95.8%; Pred. No. 0; Mismatches 127; Indels 82; Gaps 18;
Matches 4826; Conservative 0;

Qy 1 ACCCTAATACTAATCGATCCAAATATGGAAGAATTCACGTTGACTTATGATGCTGACAGC 60
Db ACCCTAATACTAATCGATCCAAATATGGAAGAATTCACGTTGACTTATGATGCTGACAGC 84
Qy 61 CCGTATGTCAAGTCGTTACAGCGGACGTTTCCACAATTTGAGATGAGCAAGCAGGCTC 120
Db CCAATATGTCAAGTCGTTACAGCGGACGTTTCCACAATTTGAGATGAGCAAGCAGGCTC 144
Qy 121 ACTGACAAATGACCAATGCGCAATGCGAGGCGTTTTCGATGCGCAACAAACCTGATTGAG 180
Db ACTGACAAATGACCAATGCGCAATGCGAGGCGTTTTCGATGCGCAACAAACCTGATTGAG 204
Qy 181 AGCGAAGTCGACCGGAGCAAGTTATCTTGACATTTGGAAGTGGCCCGCTGACATGCA 240
Db AGCGAAGTCGACCGGAGCAAGTTATCTTGACATTTGGAAGTGGCCCGCTGACATGCA 264
Qy 241 CATTCCAAATACCGCTATCATTTGATCTGCCATGATAGCGCTGAAGACCCGAGACGA 300
Db CATTCCAAATACCGCTATCATTTGATCTGCCATGATAGCGCTGAAGACCCGAGACGA 324
Qy 301 CTACAAGGCTATGAGAAAGACTTAAGAAAGTGACAT-TACGACAAAGACATAGGCTC 359
Db CTACAAGGCTATGAGAAAGACTTAAGAAAGTGACAT-TACGACAAAGACATAGGCTC 384
Qy 360 T-AAGGCGGAGACCTGCTGGAAGTCATGTCACACGAGCGACGACATCTCTGT 418
Db TNAAGGGGAGACCTGCTGGAAGTCATGTCACACGAGCGACGACATCTCTGT 444
Qy 419 GTATGACACAGACGCCAGCTGTAGTACTTTGGAAGTGA-TGCAATATACCAAGATGTG 477
Db GTATGACACAGACGCCAGCTGTAGTACTTTGGAAGTGA-TGCAAGATGTG 504
Qy 445 GTATGACACAGACGCCAGCTGTAGTACTTTGGAAGTGA-TGCAAGATGTG 504
Db TACGAGTCATGACCGACATCAATCTTACACAGGCGCTTAAAGGATTAGACATTT 564
Qy 505 TACGAGTCATGACCGACATCAATCTTACACAGGCGCTTAAAGGATTAGACATTT 564
Db TACTGATAGGCTTTGACACGACCCCTTTTATGTACAAAAACATGGCAGGTTCCACCT 624
Qy 538 TACTGATAGGCTTTGACACGACCCCTTTTATGTACAAAAACATGGCAGGTTCCACCT 597
Db TACTGATAGGCTTTGACACGACCCCTTTTATGTACAAAAACATGGCAGGTTCCACCT 624
Qy 558 ACTTTCACACGAACTGGGCTGACGAGAGATTTGGAAGCACTGAATGGGCTCGGT 657
Db ACTTTCACACGAACTGGGCTGACGAGAGATTTGGAAGCACTGAATGGGCTCGGT 684
Qy 625 ACTTTCACACGAACTGGGCTGACGAGAGATTTGGAAGCACTGAATGGGCTCGGT 684
Db AACTCAGATCTTTCAGGAGGAGGCTTGAAGAACTCTCAATCCTTGAAGAGAGGCTC 717
Qy 658 AACTCAGATCTTTCAGGAGGAGGCTTGAAGAACTCTCAATCCTTGAAGAGAGGCTC 717
Db AACTCAGATCTTTCAGGAGGAGGCTTGAAGAACTCTCAATCCTTGAAGAGAGGCTC 744
Qy 718 CAACCTACTAATAAGATCATATTCGCTGGTTTCACAACTTACACAGAGATGATCA 777
Db CAACCTACTAATAAGATCATATTCGCTGGTTTCACAACTTACACAGAGATGATCA 804
Qy 745 CAACCTACTAATAAGATCATATTCGCTGGTTTCACAACTTACACAGAGATGATCA 804
Db CTGTTACGTAGCTGGCATCTTCCAAACGTTGTCCACCTGAAGAGGAAAGCTTACCTTCA 864
Qy 805 CTGTTACGTAGCTGGCATCTTCCAAACGTTGTGTCCACCTGAAGAGGAAAGCTTACCTTCA 864
Db GGTGATGATGGGACCATTTGTGACGCTGTGAAGGATGTGATCATAAAGATTAACGATCGC 924
Qy 838 GGTGATGATGGGACCATTTGTGACGCTGTGAAGGATGTGATCATAAAGATTAACGATCGC 897
Db GGTGATGATGGGACCATTTGTGACGCTGTGAAGGATGTGATCATAAAGATTAACGATCGC 924
Qy 898 CCAGGACTATAACGTAAGTTGAGAACTTGGCGTCCACAATGCAATCGGAGGCTTCTTG 957
Db CCAGGACTATAACGTAAGTTGAGAACTTGGCGTCCACAATGCAATCGGAGGCTTCTTG 984
Qy 925 CCAGGACTATAACGTAAGTTGAGAACTTGGCGTCCACAATGCAATCGGAGGCTTCTTG 984
Db AGTTGCAAAAGTCAAGATACGCTGCGCGCGAGAGGGTTCCTTTGCTGTGTGATCGTAT 1017
Qy 958 AGTTGCAAAAGTCAAGATACGCTGCGCGCGAGAGGGTTCCTTTGCTGTGTGATCGTAT 1017
Db AGTTGCAAAAGTCAAGATACGCTGCGCGCGAGAGGGTTCCTTTGCTGTGTGATCGTAT 1044

Qy 1018 GTACAGCCACACATTTGCGATCAGATGACAGGAGTTCTGGCAATGACGTTAGTGTGAT 1077
Db GTACAGCCACACATTTGCGATCAGATGACAGGAG----- 1078
Qy 1078 GAGCGACAAAACCTATTTGGTGGGCTTCACCAAGATTTGTGCTCAATGTTAGAGCGCAA 1137
Db -----GGGCTCAACCAACGATTTGTGCTCAATGTTAGAGCGCAA 1117
Qy 1138 AGAAATCTAACAACATGCAAGAACTATCTTACCAAGTGTGCTCCAGGCGGTTTCCAGG 1197
Db AGAAATCTAACAACATGCAAGAACTATCTTACCAAGTGTGCTCCAGGCGGTTTCCAGG 1177
Qy 1118 AGAAATCTAACAACATGCAAGAACTATCTTACCAAGTGTGCTCCAGGCGGTTTCCAGG 1177
Db TGGGCGGCTGAACATCGTGTGCTGCACTTGTGAGAGAGAGAGAGTGTGCGGAGCGC 1257
Qy 1198 TGGGCGGCTGAACATCGTGTGCTGCACTTGTGAGAGAGAGAGAGTGTGCGGAGCGC 1257
Db TGGGCGGCTGAACATCGTGTGCTGCACTTGTGAGAGAGAGAGAGTGTGCGGAGCGC 1237
Qy 1258 ACTTTCATATGAGCTGTGCTGCTGGGCTTTCAAGACCCAGAAATACATCTCATCAAG 1317
Db ACTTTCATATGAGCTGTGCTGCTGGGCTTTCAAGACCCAGAAATACATCTCATCAAG 1297
Qy 1238 ACTTTCATATGAGCTGTGCTGCTGGGCTTTCAAGACCCAGAAATACATCTCATCAAG 1297
Db AAGCTGTGTAAGCAAACTGTAAGAAAGTACCTGCGCTTTTGACTCATTTGTGATTC 1357
Qy 1318 AAGCTGTGTAAGCAAACTGTAAGAAAGTACCTGCGCTTTTGACTCATTTGTGATTC 1375
Db AAGCTGTGTAAGCAAACTGTAAGAAAGTACCTGCGCTTTTGACTCATTTGTGATTC 1357
Qy 1376 C-ACGCTTACAGCCACGCGGCTCGAT-----ATGGGCTTCCGCGCT-AGGCTCAAGCT 1427
Db C-ACGCTTACAGCCACGCGGCGCTCGAATATGCGCTTTCCGCGCTTNAAGGCTCAAGCT 1417
Qy 1428 GCTGCTTGAACCAACTGTCAAACCCGACCGGCTATTTACATGCGCATGTGAGCATCT 1487
Db GCTGCTTGAACCAACTGTCAAACCCGACCGGCTATTTACATGCGCATGTGAGCATCT 1477
Qy 1418 GCTGCTTGAACCAACTGTCAAACCCGACCGGCTATTTACATGCGCATGTGAGCATCT 1477
Db GCGTGGCTTACAGCAAAAGCTGAAGAGTGGCTGACGCGGAAAGATGAGAAAGCC 1537
Qy 1488 GCGTGGCTTACAGCAAAAGCTGAAGAGTGGCTGACGCGGAAAGATGAGAAAGCC 1546
Db GCGTGGCTTACAGCAAAAGCTGAAGAGTGGCTGACGCGGAAAGATGAGAAAGCC 1537
Qy 1478 GCGTGGCTTACAGCAAAAGCTGAAGAGTGGCTGACGCGGAAAGATGAGAAAGCC 1537
Db TGCCACCTTGTGCTCCGTAATGAAATGAAAGACCGTATGAGGCAAGTATGACTTATTA 1597
Qy 1547 TGCCACCTTGTGCTCCGTAATGAAATGAAAGACCGTATGAGGCAAGTATGACTTATTA 1606
Db TGCCACCTTGTGCTCCGTAATGAAATGAAAGACCGTATGAGGCAAGTATGACTTATTA 1597
Qy 1607 TGCAAGAGCGAGAGAGAGTATGCTGAGACACC-ACGAGACACATCAGGTTGACAACT 1665
Db TGCAAGAGCGAGAGAGAGTATGCTGAGACACC-ACGAGACACATCAGGTTGACAACT 1657
Qy 1598 TGCAAGAGCGAGAGAGAGTATGCTGAGACACC-ACGAGACACATCAGGTTGACAACT 1657
Db TACCAGGCGC-AAGAGAGAGTGGGCTTGA-CGCTATCTTTCAACCCAGGCGGCTTAT-TG 1722
Qy 1666 TACCAGGCGC-AAGAGAGAGTGGGCTTGA-CGCTATCTTTCAACCCAGGCGGCTTAT-TG 1722
Db TACCAGGCTGATGAGAGAGATGGGCTTATCTTATCTTATCTTCAACCCAGGCGGCTTAT-TG 1717
Qy 1658 TACCAGGCTGATGAGAGAGATGGGCTTATCTTATCTTATCTTCAACCCAGGCGGCTTAT-TG 1717
Db TACCAGGCTGATGAGAGAGATGGGCTTATCTTATCTTATCTTCAACCCAGGCGGCTTAT-TG 1717
Qy 1723 AATAGTAAAAAACTGGGCTATCTCAACCATTTGGCGGAAACAGTACTGTATATGACTCAC 1782
Db AATAGTAAAAAACTGGGCTATCTCAACCATTTGGCGGAAACAGTACTGTATATGACTCAC 1777
Qy 1718 AATAGTAAAAAACTGGGCTATCTCAACCATTTGGCGGAAACAGTACTGTATATGACTCAC 1777
Db AATAGTAAAAAACTGGGCTATCTCAACCATTTGGCGGAAACAGTACTGTATATGACTCAC 1777
Qy 1783 AAGAGTGGGCGAGGAGATACAAAGTGCAGCATACACGCTAAGTATGTTGA-TCCAGA 1841
Db AAGAGTGGGCGAGGAGATACAAAGTGCAGCATACACGCTAAGTATGTTGA-TCCAGA 1837
Qy 1778 AAGAGTGGGCGAGGAGATACAAAGTGCAGCATACACGCTAAGTATGTTGA-TCCAGA 1837
Db AAGAGTGGGCGAGGAGATACAAAGTGCAGCATACACGCTAAGTATGTTGA-TCCAGA 1837
Qy 1842 AAGAGTGGGCGAGGAGATACAAAGTGCAGCATACACGCTAAGTATGTTGA-TCCAGA 1900
Db AAGAGTGGGCGAGGAGATACAAAGTGCAGCATACACGCTAAGTATGTTGA-TCCAGA 1897
Qy 1838 AAGAGTGGGCGAGGAGATACAAAGTGCAGCATACACGCTAAGTATGTTGA-TCCAGA 1897
Db AAGAGTGGGCGAGGAGATACAAAGTGCAGCATACACGCTAAGTATGTTGA-TCCAGA 1897
Qy 1901 TCAACGAGAGGAGTGTGTA-----CAGATACCTGCAACCAATCGCAATCA 1948
Db TCAACGAGAGGAGTGTGTA-----CAGATACCTGCAACCAATCGCAATCA 1957
Qy 1898 TCAACGAGAGGAGTGTGTA-----CAGATACCTGCAACCAATCGCAATCA 1957
Db TCAACGAGAGGAGTGTGTA-----CAGATACCTGCAACCAATCGCAATCA 1957
Qy 1949 ACGGAGAGCGCTAACAATGACAGAGATCTTAAGACTGTAAAGCTCAGACACAG 2008
Db ACGGAGAGCGCTAACAATGACAGAGATCTTAAGACTGTAAAGCTCAGACACAG 2017
Qy 1958 TAGTGAATCGCTATTAACAGAGATCTTAAGACTGTAAAGCTCAGACACAG 2017
Db TAGTGAATCGCTATTAACAGAGATCTTAAGACTGTAAAGCTCAGACACAG 2017
Qy 2009 ACTCAGAAATACGCTTGTGATTAAGAGCAAGAGTGTGTTAAAGAGAGAGAGAGGCTC 2068
Db ACTCAGAAATACGCTTGTGATTAAGAGCAAGAGTGTGTTAAAGAGAGAGAGAGGCTC 2077
Qy 2018 ACTCAGAAATACGCTTGTGATTAAGAGCAAGAGTGTGTTAAAGAGAGAGAGAGGCTC 2077
Db ACTCAGAAATACGCTTGTGATTAAGAGCAAGAGTGTGTTAAAGAGAGAGAGAGGCTC 2077
Qy 2069 CTTGTGCTTAAACGAGTATGCTGATGATCCACCATTTCAAGATTTGCGTATGAGAGTGC 2128

Db 2078 CTTTGTGCTTAATCTGATCTGATAGATCCACATTTCCAGAGTTGGTTCGAGAGTC 2137
Qy 2129 TCAAGACACGACGACGACGACCTTCAAAAGTCCCAACCATCGAGGCTATGAGGCCAG 2188
Db 2138 TCAAGACACGACGACGACGACCTTCAAAAGTCCCAACCATCGAGGCTATGAGGCCAG 2197
Qy 2189 GTTCAGGTAATCTGGAATCATCAAAAGCGCTGTGACTAAGAAAGTCTGTTGTGAGTG 2248
Db 2198 GTTCAGGTAATCTGGAATCATCAAAAGCGCTGTGACTAAGAAAGTCTGTTGTGAGTG 2257
Qy 2249 CGAAGAGGAAAACTGCGCAAAATCATAGGATGTAAAGAGATGAGAGCTATGATG 2308
Db 2258 CGAAGAGGAAAACTGCGCAAAATCATAGGATGTAAAGAGATGAGAGCTATGATG 2317
Qy 2309 TTGCTGCTAGAGCTGTGATGATGCTTCTAAATGGGGTTAAGACCCCGTTACATCTC 2368
Db 2318 TTGCTGCTAGAGCTGTGATGATGCTTCTAAATGGGGTTAAGACCCCGTTACATCTC 2377
Qy 2369 TGTACATTGATGAGCATTTGCTGCTGCATGACAGGACGCTGCTGGCATGTATGCGATG 2428
Db 2378 TGTACATTGATGAGCATTTGCTGCTGCATGACAGGACGCTGCTGGCATGTATGCGATG 2437
Qy 2429 TCAAACTTAAGAAAGTGATTTGTGCGGGGACCCAAACATGCGGCTTTTAAACATGA 2488
Db 2438 TCAAACTTAAGAAAGTGATTTGTGCGGGGACCCAAACATGCGGCTTTTAAACATGA 2497
Qy 2489 TGTGCTGAAATGATATTTTAACTGATGATGATGATGATGATGATGATGATGATGATG 2548
Db 2498 TGTGCTGAAATGATATTTTAACTGATGATGATGATGATGATGATGATGATGATGATG 2557
Qy 2549 CTAGAGAGTGACACAGACTGTATGACGCGCATGCTCCAGCGCTTCTTACGACAAAGC 2608
Db 2558 CTAGAGAGTGACACAGACTGTATGACGCGCATGCTCCAGCGCTTCTTACGACAAAGC 2617
Qy 2609 TGAAGACGTTAACCCATGTGCTGATTAATCATATAGATACCAAGGAGCACAAAGC 2668
Db 2618 TGAAGACGTTAACCCATGTGCTGATTAATCATATAGATACCAAGGAGCACAAAGC 2677
Qy 2669 CGCAAAAGATGATCTGATTTCTAACCTGTTTCAAGAGATGGGTGAACAGCTACAGATTG 2728
Db 2678 CGCAAAAGATGATCTGATTTCTAACCTGTTTCAAGAGATGGGTGAACAGCTACAGATTG 2737
Qy 2729 ACTACAAAATCAGAAAATCATGACTGCGGCTGACATCGAGAGACTTACGCGAAAGGCG 2788
Db 2738 ACTACAAAATCAGAAAATCATGACTGCGGCTGACATCGAGAGACTTACGCGAAAGGCG 2797
Qy 2789 TTTATGCTGTCAAGTACAAAGTCAAGAAATCACTCTACTCGAGACTTCTGAGACG 2848
Db 2798 TTTATGCTGTCAAGTACAAAGTCAAGAAATCACTCTACTCGAGACTTCTGAGACG 2857
Qy 2849 TGAAGCTGTACTTACACGACAGAAAAAGCATTTGTCTGAAGACGCTGCTGATC 2908
Db 2858 TGAAGCTGTACTTACACGACAGAAAAAGCATTTGTCTGAAGACGCTGCTGATC 2917
Qy 2909 CCTGATTAAGACCTTACAGCTAATATCCGGGGATTTTCACGCTTCATTGAGAGCT 2968
Db 2918 CCTGATTAAGACCTTACAGCTAATATCCGGGGATTTTCACGCTTCATTGAGAGCT 2977
Qy 2969 GGCAGCGGAAACAGACGCTTATGACGCGTTCTTGATTAAGCGCAGACAGCTGATG 3028
Db 2978 GGCAGCGGAAACAGACGCTTATGACGCGTTCTTGATTAAGCGCAGACAGCTGATG 3037
Qy 3029 TGTTCAGATTAAGGTAAAGTCTGCTGCGGAGAGCTTTAAGACCAAGCTTTGGCCACG 3088
Db 3038 TGTTCAGATTAAGGTAAAGTCTGCTGCGGAGAGCTTTAAGACCAAGCTTTGGCCACG 3097
Qy 3089 CCAACATTGTGCTGACAGACAGAGTGGGAGAGCTTGACCCATTCAAGACATGACAG 3148
Db 3098 CCAACATTGTGCTGACAGACAGAGTGGGAGAGCTTGACCCATTCAAGACATGACAG 3157
Qy 3149 CCACTCACTGAAATGGACATGAACTTTTGTGACCAAGTTCTTTGAGATGACCTGG 3208

Db 3158 CCACTCACTGAAATGGACATGAACTTTTGTGACCAAGTTCTTTGGTGAACCTGG 3217
Qy 3209 ACAGTGGGTTATTTCCGCTCCTTACCGTGCACCTTACTTACAGGATGACACTGGATA 3268
Db 3218 ACAGTGGGTTATTTCCGCTCCTTACCGTGCACCTTACTTACAGGATGACACTGGATA 3277
Qy 3269 ACTCGCAGGAGAAATGATATGGGCTTAAATAGAGGTAGCAAGAGATTGTCAAGCG 3328
Db 3278 ACTCGCAGGAGAAATGATATGGGCTTAAATAGAGGTAGCAAGAGATTGTCAAGCG 3337
Qy 3329 GATATCCGTGATCACAAAAGCGTTGACACAGGAGGATGATATATAGGATTAATA 3388
Db 3338 GATATCCGTGATCACAAAAGCGTTGACACAGGAGGATGATATATAGGATTAATA 3397
Qy 3389 CCATCAAGACATCTCTCCAAATTAATGATGCTTCAATTAATGCGGGTGGCCCACT 3448
Db 3398 CCATCAAGACATCTCTCCAAATTAATGATGCTTCAATTAATGCGGGTGGCCCACT 3457
Qy 3449 CGTTGATGCTTGAACCAAGAGACAGGATACAACTGATCACAGCGGATTCCTATCTAAGA 3508
Db 3458 CGTTGATGCTTGAACCAAGAGACAGGATACAACTGATCACAGCGGATTCCTATCTAAGA 3517
Qy 3509 TGAAGGCAATCTGTGTTGGTATCGCGATCCTATACAGATTCCAGGAGAAAGTAG 3568
Db 3518 TGAATGGCAATCTGTGTTGGTATCGCGATCCTATACAGATTCCAGGAGAAAGTAG 3577
Qy 3569 AGTCATGGGTCATGCGCACTAATACATCAGGTGATGCTGATTTGGGAATACCTA 3628
Db 3578 AGTCATGGGTCATGCGCACTAATACATCAGGTGATGCTGATTTGGGAATACCTA 3637
Qy 3629 GGCATGCGGTAATATGACATTTATCTTGTCAATGTTAGACCCCGTACAGAAACATC 3688
Db 3638 GGCATGCGGTAATATGACATTTATGTTGCAATGTTAGACCCCGTACAGAAACATC 3697
Qy 3689 ACTACAAAGTGCGAGATCACTGATTCACACAGATGCTTAACTGTAAGGCTGTCC 3748
Db 3698 ACTACAAAGTGCGAGATCACTGATTCACACAGATGCTTAACTGTAAGGCTGTCC 3757
Qy 3749 ACCACTGAACATGCGGGAACATGTCGATGATGATGATGATGATGATGATGATGATG 3808
Db 3758 ACCACTGAACATGCGGGAACATGTCGATGATGATGATGATGATGATGATGATGATG 3817
Qy 3809 CCGAATATGATCACTCGGTGACACGCTCATTTAGGTTAACCCGTCTGTCAAGCTTA 3868
Db 3818 CCGAATATGATCACTCGGTGACACGCTCATTTAGGTTAACCCGTCTGTCAAGCTTA 3877
Qy 3869 AGAACCTGCCGAAATATGAGGTTCTTCTGTTCTTGGCAAGGACAAAGGCAAC 3928
Db 3878 AGAACCTGCCGAAATATGAGGTTCTTCTGTTCTTGGCAAGGACAAAGGCAAC 3937
Qy 3929 ACAACATGACAGGACAGACTCGGTGATGCTTGAACAACATCTATCAAGGCTCAACA 3988
Db 3938 ACAACATGACAGGACAGACTCGGTGATGCTTGAACAACATCTATCAAGGCTCAACA 3997
Qy 3989 GGTACGAGGAGGAGAGCTCAGGCTACAGAGTATCAGAGGTACATTTAGCAAGAGG 4048
Db 3998 GGTACGAGGAGGAGAGCTCAGGCTACAGAGTATCAGAGGTACATTTAGCAAGAGG 4057
Qy 4049 CTGACCAAGCTATCTTATGCTGCTAATAGCAAGGCTCAACCAAGTTCCGAGTGTGG 4108
Db 4058 CTGACCAAGCTATCTTATGCTGCTAATAGCAAGGCTCAACCAAGTTCCGAGTGTGG 4117
Qy 4109 GTGACCTGATCCGAAATGCGCGCTGCTTTTATGATGACAGCAATAGCTGCGGACGG 4168
Db 4118 GTGACCTGATCCGAAATGCGCGCTGCTTTTATGATGACAGCAATAGCTGCGGACGG 4177
Qy 4169 CTAGACTGTGAAGACGACGCTCATCATACATGCTGTAGAGCCCAATTTTCTAAGA 4228
Db 4178 CTAGACTGTGAAGACGACGCTCATCATACATGCTGTAGAGCCCAATTTTCTAAGA 4237
Qy 4229 TGCAGGAAACGAGAGGCACTTAAGCTTCGAGGCTGCTTACATGATGATGATGATG 4288
Db 4238 TGCAGGAAACGAGAGGCACTTAAGCTTCGAGGCTGCTTACATGATGATGATGATG 4297

[illegible]

RESULT 2

```

US-09-454-721A-3
: Sequence 3, Application US/09454721A
: Patent No. 6296854
: GENERAL INFORMATION:
: APPLICANT: Peter Pushko
: APPLICANT: Michael D. Parker
: APPLICANT: Jonathan F. Smith
: APPLICANT: Bruce J. Crise
: TITLE OF INVENTION: Live Attenuated Venezuelan Equine Encephalitis Vaccine
: FILE REFERENCE: Army 146
: CURRENT APPLICATION NUMBER: US/09/454, 721A
: CURRENT FILING DATE: 1999-12-07
: PRIOR APPLICATION NUMBER: US 60/111,330
: PRIOR FILING DATE: 1998-12-07
: NUMBER OF SEQ ID NOS: 3
: SOFTWARE: Microsoft Word
: SEQ ID NO 3
: LENGTH: 11459
: TYPE: DNA
: ORGANISM: Venezuelan Equine Encephalitis Virus
: FEATURE:
US-09-454-721A-3

```

Query Match	42.4%	Score 2118;	DB 3;	Length 11459;
Best Local Similarity	64.5%;	Pred. No. 0;		
Matches 3219; Conservative	0;	Mismatches 1745;	Indels 24;	Gaps 3;

OY	10	ACTATATGATCAATATATGAAAGAAATTCACGTTGACTTAAGATGTGTGACAGCCGATATATC	69
Db	30	AATTAATCTACCCAAATATGAGAAAGTTCCAGTTTACATCGAGAAAGACAGCCCATTTCTCTC	89
OY	70	AAGTCGTTACAGCGGACGTTTCCAAATTTTGAGATCGAAGCAAGCAGGTCACTGACAAAT	129
Db	90	AGAGCTTTGACGCGAGGTCTCCCGAGTTTGAGTAGAAGCCAAAGAGGTCACTGATTAAT	149
OY	130	GACCAATGCCAATGCCAGAGCGTTTTCGCTATGTGGCAACAAAGCTCATTTGAGAGGAAATC	189
Db	150	GACCATGCTAATATGCAAGAGCGTTTTCGATCTTGCGCTTCAAAACGATTCGAACGAGAGTGTG	209
OY	190	GACCGGGACCAAGTTATCTTGGAACATTTGAAAGTGGCCCGGCAGACATATGCATTTCCAAAT	249
Db	210	GACCAATCCGACAGATCTCTTGACATTTGAAAGTGGCCCGCCGACAGATTAATTTCTTAAG	269
OY	250	CACCGCTATCATTTGTATCTGCGCTATGATTAAGCGCTGAAGACCCCGACAGACTAACAACGG	309
Db	270	CACAAAGTATCATTTGTATCTGTGCGATGAGATGTGCGGAAGATCCGGACAGATTTGATTAAG	329
OY	310	TATGCAAGAAAGACTTAAAGAAAGT-----GACATTACCGACAAGAACATATGCTCTTAAG	363
Db	330	TATCAACATTAAGCTGAAGAAAGAAACCTGTAAAGAAATTAATGATTAAGAAATTGGAACAGAA	389
OY	364	GCGGCAGACCTGCTGGAAGTCAATGTCAACACACAGACCAAGAGACTCAATCTCTGTATATG	423
Db	390	ATGAAGAGAGCTGCGCCGCCCTCATATAGAGGACCTCTGACAAATTGAGACATATGTGCTCTC	449
OY	424	CACACAGACGCCAGTGTAGGTACTTTGGAAGTGTAGCAGTATACCAAGATGTGTACGCA	483
Db	450	CACGACACAGAGTGTGTGCTGCTACGAAGGGCAATGCTGTTTACAGAGATGTATACGCG	509
OY	484	GTCATATGACCCGAAATCAATCTACACAGGCGCTTAAAGAGATTAGACAAATTTACTGG	543
Db	510	GTTAGACGACCGCAAGATCTCTTATACACCAAGCCAAATTAAGGAGTTAAGTTCGCTACTCGG	569
OY	544	ATAGGCTTTTGACAGACCCCTTTATGTATACAAAACATGGCAGGTTCTTACCTCACTTAC	603
Db	570	ATATAGCTTTGACACACACCCCTTTATATGTTTAAAGAACTTGGCTGAGACATATCCATCATAC	629
OY	604	AACACGAACCTGGGCTGACGAGAGATATGGAAGCAACGTAACATTTGCTCTGCTGTACTCA	663
Db	630	TCTACCAACTGGGCGGACGAAACCGTGTAAACGGCTCGTAACATATGAGCTTATGACAGCTCT	689
OY	664	GATCTTACAGAGACAGCAAGCTTGGAAACCTGCAATCTTATGGAAGAGAGGCTCCAACT	723
Db	690	GACGTTATGGAAGCGGTACAGTAAAGGAGATGTCATCTTATGAAAGAAAGTATTTAAACCA	749
OY	724	ACTATATAGATCAATATCTCGGTTGTGTTCAACATCTTACACAGAAAGTATGACTCTGTTA	783
Db	750	TCCAAACATATGTTATTTCTCTGTTGGCTGCACATCTTACACAGAAAGAGGGAATTACTG	809
OY	784	CTAGCTGGCACTTCCAAACGTGTCTCACTTGAAGAAAGTCTAACCTTACAGGTGAGA	843
Db	810	AGGAGCTGGGCACTCGCGTCTGTATTTCACTTACGTGGCAAGCAAAATTTACACATGTCTGG	869
OY	844	TGTGGGACCAATGTGCACTGTGGAAGGGTACGTCACTCAAAAGAAATTAAGATACAGCCACGA	903
Db	870	TGTAGACTATATAGTTATTTGTGGAGGGGTAGCTGTTAAGAAAGATATGATACATCAGGCG	929
OY	904	CTATACGCTAAAGTTTGAAGAACTTGGCCCTTCCACATATGATCGCGAGGGGTTTCTTGATGTTGC	963
Db	930	CTGTATGGAAGACCTTACAGGCTATGCTGTACATATGACACCGGAGGGAATTTCTGTGCTGC	989
OY	964	AAAGTCAACAGATACGCTGCGCGGCGAGAGGGTTTCTTTGCTGTGTGTACGTATGTACCA	1023
Db	990	AAAGTGCACAGACATTTGAACGGGGAGAGGGTCTCTTTTCCCGTGTACAGTATATGTCCCA	1049

OY	1024	GCACACATTTGGATCAGATGACAGGGATTCTGGCACTGACGTTAGTGTGATGACCA	1083
Db	1050	GCTACATTGTGTGACCAATATGACTGGCAATCTGGACACAGATGTGCAGCGACGACC	1109
OY	1084	CAAAAACATTGGCTTGGGGCTCAACCAAAAGGATTGTGCTCAATGTGAGACCGAAAGAA	1143
Db	1110	CAAAAACCTGTGTGGCTCAACCAAGGTATAGTGTGTCACGGTGCACCCAGAAAAC	1168
OY	1144	ACTAACACAAATGACAGAACTATCTATTAACAGTGTGTGGCCAGGCGTTTCCAGGTGGCG	1203
Db	1170	ACCAATACCATGAAAAAATTACTTTTGGCCGTAGTGGCCAGGACTTTGTAGTGGGCA	1229
OY	1204	CGTGAACTCGTGGCCGACTTGGACGACGAAAGAACTAGGGGTGCGGAGCGCACTTT	1263
Db	1230	AAGGAAATATTAAGGAAGATCAAGAAATGAAAGGCCACTAGGACCTACGAGATGACGTTA	1289
OY	1264	ACTAGGGGCTGTGTGGGCTTTCAAGCCAGAAATTCATTCATCTACAAAGGCT	1322
Db	1290	GTCATGGGGTGTGTGGCTTTTGAAGGCAAGAAATPAACCTCATTTATTAAGGCGCCG	1349
OY	1324	GGTACGCAAAACAATTAAAGAAATGACTCGTCCGCTTTGACTCATTTGTGATTTCAAGCCTT	1383
Db	1350	GATACCCAAACCATATCAAAATGAAACAGGCAATTTCACTATTGTGTCTGCCAGATTA	1409
OY	1384	ACCAAGCCACGGGCTGATATGCGCTTCCGCGTAGGCTCAAGCTGCTCTGAACCAACT	1443
Db	1410	GGCAGTAAACACATTTGGAAATCGGGCTGTGAACCAAGAAATCAGAAATGTTTAGAGGACAC	1469
OY	1444	GTCAACCCCGACCCGGCTATTACATATGCCCCATGTGAGCATCTGTGGCTTACAGCA	1503
Db	1470	AAGGAGCCGTACCTCTCAATTACGCCGACGTCACAAAGAGCTTAATGTGGCAGCCGAT	1529
OY	1504	GAAGCTGAAGAATGGCTGTCAGCGGAAAGATCAGAGAAGCCCTGCCACCTTGCTCCCT	1563
Db	1530	GAGGTTAAGAGGTGTGCTGTAAGCCGAGAGTGTGGCGACGCTTACACACTTTGGCAGCT	1589
OY	1564	GAATATGAAAAAGAGACCGTAGAGGCGAGAAATGAATCTTATTATGCAAGGCGAGACCA	1623
Db	1590	GATGTTGAGGAGCCCACTCTGGAAACCATGTGCACTTGAATGTATCAAGAGGCTGGGGCC	1649
OY	1624	GGTAGCGTGAGACACCAAGAGGACATCTAGGGTGAACAATTAACCAAGCGGAGAGAG	1683
Db	1650	GGCTCAGTGAGACACCTGTGGCTTGTAAAGTTTACAGGTAGGCTGGCGAGAGACAAG	1709
OY	1684	ATTGGGTCTTACGCTATATCTTTCACCCAGGCGGATTAATGATGAAAACTGGCGGT	1743
Db	1710	ATCGGCTCTTACGCTGTGCTTTCTCCGAGGCTGTACTCAAGATGAAAAATTATCTTCC	1768
OY	1744	ATCCACCCATTGGCGGAAACAAGTACTGTATATGACTCAAAAAGTAGGCGACGAGATAC	1803
Db	1770	ATCCACCTCTCGCTGAACAAGTACATAGATTAACACACTGTGGCCGAAAAAGGGGTTAT	1829
OY	1804	AAAGTCAGCCATATACCAAGGTATGTCAATTGTACCAAGAAAGGACGGCGCTCCGTTCAA	1863
Db	1830	GCCGTGGAACCATATCAATGTATTAATATGTGTGTCAGAGGACATGCAATATCCGTCCAG	1889
OY	1864	GACTTCCAGCAATTAGTGTAGAGCGCTTCAGATCGTTTCAACGAAGGAGAGTGTGTAAAC	1923
Db	1890	GACTTTCAGGCTCTGATGTAAATGTCCACCATTTGTGTCAACGAACGTGAAGTTCGTAAAC	1949
OY	1924	AGATACCTGCACCATGCAATCGCAATCAACGAGAGAGCGCTTAAACACTGACGAAGATCTAT	1983
Db	1950	AGGTACTGTGACCAATATTTGCCACATGGAAGAGCGCTGAACACTGATGAAGAAATATAC	2009
OY	1984	AAAGATGTAAAGATCTCAGGACACAGACTTCAGAAATGCTCTTGCGATTTGACGCACGAAG	2043
Db	2010	AAAACTGTCAAGCCAGGAGACACAGACGGCGAATCTGTACGATGACAGAGAAACAG	2069
OY	2044	TGTGTTAAGCGAGACGACAGGTCCCTTGTGCCCTAACCGGAGATCTGTGATTCACACA	2103
Db	2070	TGCTGTAAAGAAATCTGTACTGTGGCTTACGGGCTTCAAGGCGAAGCTGTGTGATCTTCC	2129
OY	2104	TTTACGAGTTTGCATGAGAGTCTCAAGACACGACGACGACCTTCACAAAGTCCCA	2163

Db	2130	TTCCATGAAATTCGCTTACGAGAGCTCTTGAAACACACACAGCGCTCTTACCAAGTATCCA	2188
Qy	2164	ACCAATCGAGAGCTATGAGATGCGACAGGTTCAAGTAAATCTGMAATCATCAAAAAGCGCTGTG	2222
Db	2190	ACCAATAGGGGTGTATGTGGCTGCCAGSATCAGGCAAGCTTGGCATATTTAAAGCCAGTC	2244
Qy	2224	ACTAAGAAAGATCTGGTTGTGATGTCGAGAAAGGAAAACTGCGCAAAATCATCAGGAGT	2282
Db	2250	ACCAAAAAAGATCTAGTGTGTGAGCGCCAAAGAAAGAAAATCTGTGCAGAAATTTAAGGGAC	2300
Qy	2284	GTAAGSAGATGACAGCTATGATGTCTGTCTAGSACTGTTCGATTCAGTGTCTTAAAT	2344
Db	2310	GTCAAGAAAAAGAAAGGGCTGAGCGTCAATGCAAGACCTGTGACTCAGTGTCTTTGAAT	2366
Qy	2344	GGGGTTAAGCAACCCCGTTAAACCTCTGACATTTGATGAGGCAATTTGCCGCAATCGAGG	2400
Db	2370	GGATGCCAAACACCCGTAGAGACCTGTATATTGACGAAAGCTTTGTCTGTATCAGGCT	2422
Qy	2404	ACGCTGTGAGCACTGATTTGCCATGCTCAAACTAAGAAAGTGTATTTGTGCGGGACCCA	2466
Db	2430	ACTCTCAGAGGGCTATAGCCATTATAGACTTAAGAAAGGCAAGTCTCTGGGGGATCCC	2488
Qy	2464	AAACATGCGGCTCTTTTAACATGATGTGCTGAAAGTACATTTTAACCATGACATATGC	2522
Db	2490	AAACAGTCCGGTTTTTTTAACTATGATGTGCTGAAAGTCAATTTAACACACAGATTTTGC	2544
Qy	2524	ACTGAAGTATACAAATTAAGATCTCTGAGAGGTGACACAGACTGTAAACGCGCATTCGT	2588
Db	2550	ACACAGTCTTCCACAAAAGCATCTCTCGGTGCACTTAATCTGTGACTTTCGGTCTGC	2600
Qy	2584	TTCAAGCTCTTCTTACGACAGAGSAGATGAAGACGTTTAAACCATGTGCTATTTAAATCATC	2644
Db	2610	TCAACTCTGTTTACGACAAAAATAGAAAGACAGATTCGAAAGAGACTTAAGTTGTG	2666
Qy	2644	ATGATATCCACAGGAGCAACAAGCCGACAAAGATGATCTGATTTCACTCTTTTACA	2700
Db	2670	ATTGACACTACCGGACAGTACCAAACTTAAGACGACGATCTCATTTCTCACTTTTACAGA	2722
Qy	2704	GGATGGGAGAAACAGCTACAGATTGACTACAAATAACGAAATGATCACTGCGGCTGCA	2766
Db	2730	GGGTGGGTGAAGAGCTTGCAAATGATTTACAAAGGCAAGAAATATATACGGCAGCTCC	2788
Qy	2764	TCGCAAGACTTACGCGGAGAAAGCGCTTATGTCTGTCAAGTCAAAAGTCAACGAGAAATCCA	2822
Db	2790	TCTCAAGGGCTGACCCCGTAAAGGTGTATATCCGTTACGATCAAGGTGAATGAATCTCT	2844
Qy	2824	CTTACTCGACAGCTTCTGAGCAAGTGAACGTGTATCTTACACGACAGAAAAAGCAATT	2888
Db	2850	CTGTACGACCACTCATGAAACATGTGAACGTCTCTACGACCCGACAGGAGACGCGATC	2900
Qy	2884	GTCTGGAAGACGCTAGCTGTGTATTCCTGTGATTAAGAGACTTAACGCTAAATATCCCGG	2944
Db	2910	GTGTGAAAAACAATGACCGGCGACCCCAATGAATAAAAACAATGACCTGCAAGTACCTTGG	2966
Qy	2944	GATTTCACGGCTTCAATTGAGCAAGCTGCGAGCGGCAACAGACGCTTATGTGACAGCGCTT	3000
Db	2970	AATTTCACTGCCACGATAGAGAGGTGGCAGACAGACATGATGCCATATAGAGCAATC	3022
Qy	3004	CTTGTAAACCGCAGACAGCTGATGTGTTCAGAAATGAAGTGAACGTCTGCGGCGAAG	3066
Db	3030	TTGGAGAGACCGGAGACCTTACGAGAGCTTTCAGAAATAGGCAAAAGTGTGTGGGCGAAG	3088
Qy	3064	GCTTTAGAGCCAGCTTTGGCCACGCGCAAACTTGTGTGACAGAGACACAGTGGGAGACG	3122
Db	3090	GCTTTAGTGGCGGTGTCTGAAGACCGGTGGACATGACATGAAACAATGGAACACT	3144
Qy	3124	TTGCACCCATTCAAGCATGACAGAGCGTACTCACTGAAATGGCACTGAACCTTCTTTTGC	3188
Db	3150	GTGATTTATTTTGAACGGACAAAGCTCACTCAGAGAGATATGTTGAACCAACTATATGC	3200
Qy	3184	ACCAAGTTCCTTGAAGTGAACCTGAGACAGTGGATTATTTTCCGCTCTTACCGTGCACATT	3242

3210 GTGAGTCTTTGGACTGATCTGAGCTCCGGTCTATTTTTCGACCCACCTGTTCCGTTA 3269
3244 ACTTACAGGATCGACCTGGGATTAATCCGCGAGGAAAGAACATGATGGCTTAATAGA 3303
3270 TCCATTGAGATTAATCACTGGGATTAATCCCGCTGCTTAATGATGGGCTAAATPAA 3329
3304 GAGGTACCAAGAGATTGTCCGCGGATATCCGTGATCAACAAAGGGGTGACACAGGC 3363
3330 GAATGGTCCGTGACTCTCTCGAGGATACCAACATGCTCGGGGAGTTGCCATGGA 3389
3364 AGGTAGCTGATATAGGAATTAATCACTCAAGAGATCTCTCAACATTAATATGCTT 3423
3390 AGAGTCTATGACATGAACACTGTGATCACTGGCAATTAATGATCCGCGATTAACCT 3449
3424 CCATTAAATGGCGGCTGCCCCACTCGTTGATGCTTGAACCAAGACAGGATCAACT 3483
3450 CCGTAAACAGAAACTGCTCTGATGCTTATGCTTCCACCAATTAAGAACCCACAGAGT 3509
3484 GATCACAGCGGATTCCTATCTAAGATGAAGGCAAAATCTGTGTTGATCGCGATCCT 3543
3510 GACTTTCTTCTATCTGTCAGAAATGGAAGGCAAGAACTGCTGTGTTGCGGGAAAG 3569
3544 ATCAGACTTCCAGGAAAGAAAGTGAAGTCAATGGGCTCATTGCCATTAATACATCAG 3603
3570 TTGTCCTCCAGGCAAAATGTTGATGCTGTGATCAGACCGGCTGAGGCTACCTTGA 3629
3604 TGTGATCTCATTTTGGGAATCTAGCAATGCTGTTAAATGATTAATCTTTGTCAAT 3663
3630 GCTGCGTGAATTAAGGATCCAGGTATGTCCTCAATTAATGATTAATTTTGTAT 3689
3664 GTTACAGCCCCGTAACAGAACTATCACTCAACAGTGCAGAGTCACTATCCACAC 3723
3690 GTGAGAACCCATTAATTAATCACTATCAGAGTGTGAAGACATGATTAATAGCTT 3749
3724 AGCATGCTAAGCTGTAAAGCTGTCCACCACTGAACACTGCGGAAATGTGTGCTATA 3783
3750 AGCATGTTGACCAAGAAAGCTTGTCTGATCTGATCCCGGGAACCTGTGTGACGATA 3809
3784 GGGATATGGCTTGTGATTCGCGCAACGAAATATCACTGGGTTGGACGCTCATTT 3843
3810 GGTATGTTAGTGTGATGAGGCGCAAGAACATCATTTGTTGATTAAGCGGAGTTC 3869
3844 AGTTTACCCGCTGTCTGTCACTAAGAACACTGCGGAAATTAAGTGTCTTCTG 3903
3870 AAGTTTCCCGGATATGCAACCGAAATCTCACTGTAAGAGAGGAAGTTCTGTGTTGA 3929
3904 TTCTTCCGCAAGAACGCAACCAACATCACTGACGACAGACTGCTGTGTAGTCTT 3963
3930 TTCAATGGGTATGATCGCAAGGCGGCTGACGCAATCTTCAAGCTTTCATCAACCTTG 3989
3964 GACAACTATCTAAGAGGTCAACAGATAGAGGAGGAGGAGCTCCAGCGTACAGATG 4023
3990 ACCAACTTAATTAAGGTTCCAGACTCCACAGAGCGGATATGACCTCTCATATCATGTG 4049
4024 ATCAGAGGTGACATTAAGCAAGAGCGCTGACCAAGCTATCTTAAATGCTGTAATAGCAA 4083
4050 GTGGAGGGGATATTTGCCAGGCGCACGAGAGATTAATTAATGCTGCTAAACAGCAA 4109
4084 GGTCAACCAAGTTCGCGAGTGTGCGGTCACTGTACGAAATGCGCGCTCTTTGAT 4143
4110 GGAACAACCTGGCGAGGGGTGTGCGAGCGCTGTATAAGAAAGTTCCCGAAAGCTTCAT 4169
4144 AGAAGCAATATGCTGTGCGGACCGGTGATCTTGTGAAGCAGAAC--GCTATCTA 4200
4170 TTACAGCGGATCGAAGTGAAGAAAGCGGACTGTGCAAGGTGACGCTTAACATATCAT 4229
4201 CATGCTGAGGACCAATTTTCTAAGATGCGGAAACGAGGGGACCTTAAGCTGCA 4260
4230 CATGCGTGAAGACCAATCTTAACAAAGTTTCCGAGGTGAAGGTGAACAAACAGTTGCA 4289
4261 GCTGCTTACATGATGATGCTCATGCTCAACGCTGAGCGGATTAACAAAATATCAGTA 4320
4290 GAGGCTTATGATGATGCTCATGATGATGTCAACGATTAACAAATCAAGCTAGTACGAT 4349

4321 CCGCTACTGTCACCCGCGCATCTATCTGTGTGCAAAAGATGAGATGATCATCTGAT 4380
4350 CCACTGTGTTCACCGGCACTTTTCCGGAACAAAGATGATCAACCCATCATTTGAA 4409
4381 CACTGTTCACGCTTTTGAACACTACAGATGCGGATGCAATATATGCTTGGATPAA 4440
4410 CATTTGCTGACAGTTTGAACACCACTGATGAGATGATGATCAATATCTCAGGAGCAAG 4469
4441 CAATGGAGACCAAGATTAATGAGGCTTACCGGCAAAAGAACCTGCAATTTGGA- 4499
4470 AAATGGGAATGACTCTCAAGGAAGCGTGGCTAGGAGAAAGCAGTGGAGAGATGTC 4529
4500 -----TGATGACAAAGCCAGTGAACATTAAGTGTGTCAGGCTCAACCCAAAC 4545
4530 ATATCCGACGACTTTTATGATGACGAACCTGATGACAGAGCTGTGAGAGGTGATCCGAG 4589
4546 AGCTTTTGGCAGGACAGCAGGTTACTCCGTCAATAGGGCAAGTTGATTCATACCTG 4605
4590 AGTTCTTTGGCTGGAAGGAAGGGCTACAGCAACAGCATGGAACAACTTTCTCATATTTG 4649
4606 GAAGTTCACGATTCATCAACCGGCAAGACATTTCCGAATTCATGCAATGTGCCCC 4665
4650 GAAGGACCAAGTTTCAACAGCGGCGCAAGATATGACAAATTAATGCTATGTGCCCC 4709
4666 AACAAATCTGAGGCTAATGACAGATTTGCTTGTATCACTCGGGGAGAGATGTCCAGC 4725
4710 GTTGCAACGAGGCGCAATGAGAGTATGATGATGATCTCTCGAGAAAGATGATGAGCT 4769
4726 ATCCGCTCAAAATGCGCAGTGAAGAGTCAAGAGCGTCTGCTCACTCACTCACTTCCA 4785
4770 ATTAGTTCGAATATGCGGCTGGAAGAGTCCGAAAGCTTCCACACCACTTACGACCTGCT 4829
4786 TGCTGTGTATTAAGCTATGAGGCTGAGCGCTATACAGGTTGCCCTCTGCGAAGAAA 4845
4830 TGCTTGTGATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4889
4846 GAACAGTTCGCGGATGCTATCACTCTGTTGCGGAAGTACAGATACAGAGGCTGAG 4905
4890 GAACAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4949
4906 AAGCTACATGACAGCAAAAGCTGCTTGTTCAGGCGCTGATCAACCGGCTGTACACCCC 4965
4950 AAGATCCAAATGCTCCAGCTATATGTTCTACACGAAAGTGTGCTGTATATTCATCA 5009
4966 AGGAAGTA 4973
5010 AGGAAGTA 5017

RESULT 3
US-08-991-840A-2
; Sequence 2, Application US/08991840A
; Patent No. 6261570
; GENERAL INFORMATION:
; APPLICANT: Michael D. Parker
; APPLICANT: Jonathan F. Smith
; APPLICANT: Bruce Crise
; APPLICANT: Mark Steve Oberste
; APPLICANT: Shannon Schumura
; TITLE OF INVENTION: Live Attenuated Virus Vaccines for Eastern Equine Encephalitis
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Charles H. Harris
; STREET: USA MWC - MWC-0A
; CITY: FORT DETRICK, FREDERICK
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.5

SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991.840A
FILING DATE: December 16, 1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Provisional Application 60/047162,
FILING DATE: May 20, 1997
APPLICATION NUMBER: Provisional Application 60/053,652
FILING DATE: July 24, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Charles H. Harris
REGISTRATION NUMBER: 34,616
REFERENCE/DOCKET NUMBER: 003/058/SAP RIID 96-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 619-2065
TELEFAX: (301) 619-5034
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 11464 base pairs
TYPE: Nucleic acid
STRANDEDNESS: Double
TOPOLOGY: Unknown
US-08-991-840A-2

Query Match 41.5%; Score 2074.8; DB 3; Length 11464;
Best Local Similarity 64.1%; Pred No. 0;

Matches 3202; Conservative 0; Mismatches 1772; Indels 24; Gaps 4;

QY 1 ACCCTCAAACTAATCGATCCAAATATGGAAGAATTCAAGTTAGTACTGATGCTGACAGC 60
DB 20 AGCCCAAAACAATACTACCCCAAAATGGGAAGTTACGTTGATCATGAGAGAAATATGT 79
QY 61 CCGTATGTCAGTCTGTACAGCGACGTTTCCACATTTGAGATGGAAGCAAGCAGGTC 120
DB 80 CCGTCCCTCAGAGCATTAACAGGAGCTTCCGAGTTTGAGTGAAGCCCAAGCAGGTC 139
QY 121 ACTGACATGACCATGCGCAATGCGAGAGGTTTGGCATGTGCAACAAAGCTCATTGAG 180
DB 140 ACAGATATGACATGCTAAGCCAGAGGTTTGGCATTTGGCATTCGAATTTGATCGAG 199
QY 181 AGCGAAGTCGACCGGGAACCAAGTATCTTGAACATTTGAAGTGCSCCGTCAGACATCA 240
DB 200 ACGGAGTGGAAACCATCCGATTCGATCTAGACATTTGGAAGTGCSCCGTCAGAAATG 259
QY 241 CATTCCATACCGCTATCATTTGTATCTGCCCTATGATGAGCGCTGGAAGCCCGACAGA 300
DB 260 TATTCAGAGCATTAAGTACCATTTGATCTGTCCGATGAATGTGCAGAAATCCGACAGA 319
QY 301 CTACAAAGGTATGAGAAAGACTTAAGAAAAGT-----GACATTACCGACAAAGAACATA 354
DB 320 CTGTTTAAGTATGACGCCAAGCTGAAGAAACCTGTAAGAGATTACAGATTAAGAGACTG 379
QY 355 GCCTTAAGGGCGGAGACCTGCTGGAAGTCATGTCAACACGACGACAGAGATCCATCT 414
DB 380 GACAAAGAGATGAAGAGAGCTTGCAGAGTCATGACGACCCGATCTGGAACCTGAACG 439
QY 415 CTGTGTATGACACAGAGCCACGTTGAGTACTTTGGAAGTGAAGTGAAGTATCAAGAT 474
DB 440 ATTTCCTTCAAGAGATGAACCTGTGATTTGAGGTCGAAGTCGAGTATACAGAT 499
QY 475 GTGTACGAGTCATGACACCGGACATCAATCTACACGAGCGCTTAAGAGATTGAGACA 534
DB 500 GTGTACGAGTGTGAGACCGGACCGAGCTTTACATCGGCAACAAAGGGTTCAGATC 559
QY 535 ATTACTGATAGCTTTGACACGACCCCTTTATGTATCAAAAAATGAGAGTTCTTAC 594
DB 560 GCCTATTTGATGATTTGACACATACCTTTTATGTTTAAGAACCTGGCTGAGACATAT 619
QY 595 CCTACTTACACAGAACTGGGCTGACGAGAGATTTGGAAGCAACGTAACTTGGCTC 654
DB 620 CCCTCTATTGTCACCACTGGGCGGACGAGACGAGTGTAAACGCTCGTAAATTAAGCTTG 679

QY 655 GGTAACTCAGATCTTACAGAGAGAGGCTTGAAAACTCTCAATCTTTAGNAGNAGAGG 714
DB 680 TGCAGCTCCGATGTGATGAGAGGTCACGTAGAGGATGTCCATCTCAGAAAGAAATTT 739
QY 715 CTCAACTTAAATATGATCATTTCTGGTTGGTTCAACATTTACACAGAGATAGA 774
DB 740 TTAAAACCATCAATTAAGTCTTGTCTGTAGATCTTACATCTACAGAGAAAGCA 799
QY 775 TCAGTTACGATGAGTCATCTTCCAAACGTTTCCACTGAAGAAAGCTTAATCTTC 834
DB 800 GACTTACTAAGAGTTGGACCTTACCTTCCTGTTTTCACCTACGTTGAGAGAAATTA 859
QY 835 ACAGTATAGTGGAGACCATTTGTACGTGTGAAGGTAAGTATCAATAAAGATAAGATC 894
DB 860 ACTGTGTGTGAGACTATATGTTAGTGGACGGGTGATGTGTCAAAAGGATACCTATT 919
QY 895 AGCCAGAGACTATACGTTAAAGTTGAAACTTGGCTCCACATATGATCGGAGGTTTC 954
DB 920 AGTCAGGTCTGTACGGGAAACCGTCCGGCTACGCTGCACAGATGCAATCGGAGGATTC 979
QY 955 TTGAGTTGCAAGTACAGATACGCTGCGGCGGAGAGGTTTCTTTGCTGTGTACG 1014
DB 980 TTGTCTGCAAGGTGACGGAACGCTTGAACGAGAGAGGCTCTTTTCCCGTATGACG 1039
QY 1015 TATGTACAGCCACACTTGGCATCAGATGACAGGATTTGCGCACTGACGTTAGTGTG 1074
DB 1040 TACGTCCAGCACAATTTGTGGACCAATGACAGGATTTCTGGCAACAGATGTCACTGCA 1099
QY 1075 GATGACGCAAAAATCTATTTGGTGTGCTCAACCAAGATTTGCTCAATGTATGACG 1134
DB 1100 GACGACGCGCAAAATTTCTGTTGGCTCAACACGCAATATGCTAAATGCGCGCACT 1159
QY 1135 CAAGAATTAATACACATGACAACTATCTATTTACAGTGTGCGCCAGCGCTTCC 1194
DB 1160 CAAGAATTAATACACATGACAACTATTTATTTTACCGTGTGCGCAAGACATTTCC 1219
QY 1195 AGTGTGGCGCTGAACATCTGTGCGCATTTGACGACGAGAAAGAACTAGGGGTGCGGAG 1254
DB 1220 AGATGGCTTAAGATGACACAGAAATGACAAAGATGAAGAACATTTGGCTTAGGAGC 1279
QY 1255 CGCACTTACTATGCGCTGCTGCTGGCTTTCAAGACCCAGAAATCAATCTCATC 1314
DB 1280 CGCAGTTGTATATGGGGTGTGTTGGGCTTACAGAAACCAAGATTAATCACTAGTAC 1339
QY 1315 AAGAAGCTGTGATGCGCAAACTTAAGAAATCTGCGCTTTGACTCATTTGTGATT 1374
DB 1340 AAAGCAACGACACCCAAACGATCATCAAGGTAAACAGCGATTTCCACTTTTCTGTG 1399
QY 1375 CCAGCCTTACAGCAGCGGCTGATATGAGGCTTCCGCGTAGGCTCAAGCTGTGCTT 1434
DB 1400 CCGAATTTGAAAGCACAACCTTGAAGTTGGCTGAGGACCAAGATCAAGAACTACTG 1459
QY 1435 GAACCACTGTCAAAACCGGACCGGCTATTCAATGCGCGATGTGAGCATCTGCGTGC 1494
DB 1460 GAGGAACCTGTGACAGACCACTGATTTACCGCGGACGATACAGAAAGCCAAAGAC 1519
QY 1495 TTAAGCAAGAACTGAAAGAGTGGCTGACCGGAAGAGATCAAGAAAGCCCTGCCACCC 1554
DB 1520 GCGGCGATGTAGGCTTAAGAAAGTCAAGGAGCCGGAAGGCTCAAGGAGCATTTACACCG 1579
QY 1555 TTGCTCCCTGAATAGAAAAAGAGCCGTAGAGGAGAAAGTACCTCATTAATGCAAG 1614
DB 1580 CTGTCTGCGATGTGAGAAACCTGCACTGAGAGGCGGACGTTGACTTAATGCTGAGAG 1639
QY 1615 GCAAGAGCAGGTACGTTGAGACACACAGAGACACATCAAGGTGACAAATTAACCAAGC 1674
DB 1640 GCGGAGCAGGATCTGTGAAACCCCGAGGGGCTCATCAAGTACACAGTTATGACAGA 1699
QY 1675 GAAGAGAAATTTGGCTTTAAGCTTATCTTACCCAGCGCGGTATTAATGAGTAA 1734
DB 1700 GAAGCAAAATTTGGCTTTAATGCGGTGCTTCCCAACAGAGTATCGCAAGTAA 1759
QY 1735 CTGGGCTATTCACCCATTTGGCGGAACAGTACTGTGATGACTCAAAAGGTAGGCA 1794

Db 1760 CTGACGTGATCCACCCGCTTGACAGAGCAAGTAATGTATACACACTCTGGAAGAG 1819
Qy 1795 GGGAGATACAAAGTCGAGCCATACCGGTAAAGTCAATTGTACCAAGAGGACGGCGTC 1854
Db 1820 GGTGATATACGAGTTGAGCCCTTACACGGAAGGTGTAGTACCCGAAGGGCAAGCTATA 1879
Qy 1855 CCTGTCAAGACTTCGAGCAATGTAGTAGAGCGCTACGATGTTTTCAAGAGGGAG 1914
Db 1880 CCTGTCAAGACTTCGAGCAATGTAGTAGAGCGCCCAATTTGTGTACAAAGAGCTGAG 1939
Qy 1915 TTGCTAAACAGATACCTGACACCAATGCAATCAACGAGGAGCGCTAAACCTGACGA 1974
Db 1940 TTGCTAAATAGGTATTTACACCAATGCGACGACGGTGGAGCTTGAAACACAGACGAG 1999
Qy 1975 GAGTACTATAGACTGTAAAGACTCAAGACACAGACTCAAGTAATAGCTTTGATTTGAC 2034
Db 2000 GAGTATTTACAGAGTGTGAAACCTAGTGAACATAGAGGTGAGTACTTATATGACATTGAC 2059
Qy 2035 GCACGAAAGTGTGTTAAGCGAAGACGACAGTCCCTTGCTAACCGGTGATCTGGTA 2094
Db 2060 AAGAAACAATGTGTGAAAGAAAGAACTGTTACGGGACTAGTTTGAACGAGGAATTGTTC 2119
Qy 2095 GATCCACCATTTCAAGATTGCGTACGAGACTCAAGACACGACGACGACACTTAC 2154
Db 2120 GATCCCGCTTCATGATTTGCAATGAAAGCTTGAAGACACGTCGCGCCGCACTTAT 2179
Qy 2155 AAAATCCCAACCATCGAGTCTATGAGTGCAGGTTCAAGTAATCTGMAATCATCAA 2214
Db 2180 CAAATACCAACTAATAGGTGTATGCGTTCCGCGTCCGGAAAGTCTGGAATTAATAG 2239
Qy 2215 AGCGCTGACTAAGAAAGATCTGTGTGAGTGCAGAAAGAAACTGCGAGAAATC 2274
Db 2240 AGCCCAATGCAAAAGAGAAAGTGTGTAGTGTGCTAAGAAAGAACTGCGCCGAGATA 2299
Qy 2275 ATCAGGATGTAGAGAGATGAGACGTATGATGTTGCTGCTAGGACTGTGATTCAGTG 2334
Db 2300 ATAAGGACGTCAAGAGATGAAGGCTAGACGTCAATGCCCCGAGCGGTGACTCAGTG 2359
Qy 2335 CTCTTAATGGGGTTAAGCAACCCCGTTAACACTCTGTACATTTGATGAGGCAATTCCTGC 2394
Db 2360 CTGTTAATGTAGTCAAGCAACCCCGTGAAGCTTTGTACATGTATGAAGCCCTTTCATGTC 2419
Qy 2395 CATCAGGAGCGTGTGCACTGATTCGATGCTCAAACTTAAGAAAGTGTATGTC 2454
Db 2420 CAGCCGCGCACTCTCAGGCGCTTAATAGCACTTATAGCCCAAGAAAGCACTGTATGT 2479
Qy 2455 GGGGACCCAAACCAATGCGGCTTCTTAACATGATGTGCTGAAGTACATTTTACCAT 2514
Db 2480 GGTGACCCAAAGCAATGTGCGCTTCTCAATGATGTGCTGAAGTGCATTTCAACCAT 2539
Qy 2515 GACATATGCACTGAAGTACATTAAGTCTCTGAGAGGTGCACACAGACTGTACC 2574
Db 2540 GAAATATGCACTCAAGTTTCTCAAAAGCAATATCTGCAAGTGTACCAATCTAGTACG 2599
Qy 2575 GCATCTCTCCACGCTCTTCTACGACAGAGGATGAAGCGTTAAACCATGCTGAT 2634
Db 2600 TCGGTAGTGTCCACACTGTTTATGACAAAGAAATGAGAACTACTAACCCCAAGGATTC 2659
Qy 2635 AAAATCATATGATACCAAGGACCAAAAGCGCAAAAGATGATCTGATTTCAAC 2694
Db 2660 AAAATCAAAATTTGACAAACAGAAATGACTAAACCAAAAAAGATGACTTATCTCA 2719
Qy 2695 TGTTCAGAGATGTGGTGAACAGCTACAGTATGATCTACAAATCATGAAATCATGACT 2754
Db 2720 TGTTTTAGGGATGGTTAAGCAACTGCMAATATACCTAAAGGAAATGAAATATATGCC 2779
Qy 2755 GCGGCTCATCGCAAGACTTACGCGGAAAGCGCTTATGCTGAGGTACAAAGTCAAC 2814
Db 2780 GCGGCTGCTCACAGGATTTACCGGAAAGGTGTCTATGCACTTAAAGTATCAAGTTTAC 2839
Qy 2815 GAGATCACTCTACGACAGCTTCTGAGCAGCTGAACGTGTTACTTACACGACAGAA 2874

Db 2840 GAGAACCAATTTATGACGACCCACCTCAGACATGTAAATGTGTGTAGCCGCGAA 2899
Qy 2875 AAAAGCATGTCTGGAAGAGCTAGCTGTGATCCCTGGAATTAAGACATTTACGTTAA 2934
Db 2900 GACAAAGTTGTGTGAAGACTCTTTCAGAGGAGACCCGTGTATTAAGACCTTGACCGCAAG 2959
Qy 2935 TATCCCGGAGATTTCAAGGCTTCAATGACACTGAGCGAGCGAAACGACGCAATTTAT 2994
Db 2960 TACCCGGAATTTCACTGCAACAATGGAAGATGGAAGGACGAACATGATGCATATG 3019
Qy 2995 GCACGCTTCTTATTAACCGCAGACAGCTGATGTGTTCCAGAAATTAAGTGAAGTCTGC 3054
Db 3020 AGACACATCTGTGAGAAACCGGATCCCAAGATGTCTTCCAAATTAAGTGAATGTTTGC 3079
Qy 3055 TGGGCGAAGCTTTAGGCGCAGTCTTGCCACGCGCAACATGTCGACGAGACAG 3114
Db 3080 TGGGCAAGGACACTTGTACTGTGTGTTAAGACAGCCGGGATGATTTGACCAAGACAG 3139
Qy 3115 TGGGAGAGCTTGCACCCATTCAGACATGACAGAGCTACTACCTGAAATGCACTGAAC 3174
Db 3140 TGGAAACAGTGTATTTCTTCAAGAGAGATTAAGGCCCACTCAGCTGAGATTTGCTGAT 3199
Qy 3175 TTTCTTGCACAGGTTCTTTGAGTAGACCTGACAGTGGGTTATTTCCGCTTACC 3234
Db 3200 CAGCTGTGCGGATTTCTTGTGCTAGTACTTATGATTTCTGTTTTCGCCCCACA 3259
Qy 3235 GTGCACTTCTTACAGGAGTACGACATGGGATTAATCGCGAGGAAGAAATGATGG 3294
Db 3260 GTTCACCTCTCATTAAGAAACAACATTTGGAGAACACTACCGTCAACCAACATGTAGGG 3319
Qy 3295 CTTAATAGAGGTAGCAAGAAAGATTGTCAAGCGATATCCGTGATCAGAAAAGCGTT 3354
Db 3320 TTGAATCAAGAGTGTGACAGCAACTATCAAGCAGATACCTCAATTAACACGTCGGTG 3379
Qy 3355 GACACAGCAGGTTAGCTGATTAATGAGAAATTAATCAATCAAGGACTACTCCAACTT 3414
Db 3380 ACTACTGGAGAGATGACATGAACACCGGACTTGTGCAATTTGATTCGCGCATTA 3439
Qy 3415 AATGTGTTCCATTAATTCGCGGTTGCCCACTCGTGTATTCGTGACCAAAAGGACAG 3474
Db 3440 AATTTATACCGGGAACCGTGTCTACACATGCTCTCGTGAACCAATGCTGATCAT 3499
Qy 3475 GGTACAACTGATCAACAGCGATTCCTATCTAAGTGAAGGCAAAATCTGTGTGATC 3534
Db 3500 CTTCCAGTATTTTTCGCGCTTTGTCAAGTATTAAGGCAAGAACGTCCTAGTATG 3559
Qy 3535 GCGGATCTTATCAGACTTCAAGGAGAAAGTATGATCTCATGGTCCATTCCTAAT 3594
Db 3560 GGTGAGAAAGTGAATTTTCAAGTGAAGACGCTGATGCTTATCTGAAACACCTGATTC 3619
Qy 3595 ACCATCAGGTGATCTCGATTTGGGAATCTTAGCCATGTCGGTAAATATGACATATC 3654
Db 3620 ACTTTAGGCGCGCTAGATCTAGGCATACCAATGACTACCGAAGTACGATATGTC 3679
Qy 3655 TTTGTCAATGTAGAACCCCGTACAGAAACCATCATCAACAGTGCAGATCAGCT 3714
Db 3680 TTGCTAATGTAAAGAACAGTACCGCTACCAACAAGACAGTGTGAGAACAGGCC 3739
Qy 3715 ATCCACCAACAGATGCTAACGTGTAGGCTGTCCACCACTGAACAATGCGGAACTGT 3774
Db 3740 ATTAAGTTGAGACATGTTGACCAAGAGGCTGTGCTGCACTGAAACCCGAGAGAACCTGT 3799
Qy 3775 GTGCTATTAAGGTATGGGCTTGTGATTCGCGCAACGAGAAATATCATCTGGGTGGCA 3834
Db 3800 GTAGCATGTGTATCGGCTATGCGGACCGGCGAGTAAAGCAATCAATAGGTGAGTGTCT 3859
Qy 3835 CGCTATTTAGTTTACCCGCTGTCTGACAGCTTAAGAACACTCCGAAATTAATGAGTT 3894
Db 3860 CCGGAGTTCAAGTTCTGAGGATATGCAAACCGAAGTGTCTTAAGAGAGAGACGAAAGTG 3919
Qy 3895 CTCTTCGTGTTCTTCCGCAAGGACCAACGGAACCAACATGACAGACAGACTCGGT 3954
Db 3920 CTATTTGTCTTACATGGGTGATCGTAAACCGAAACCATTAACCATTAAGCTCTCC 3979

[illegible]

```

      4 RESULT 4
      5 US-07-920-281C-1
      6 ; Sequence 1, Application US/07920281C
      7 Patent No. 5739026
      8
      9 GENERAL INFORMATION:
     10 APPLICANT: Garoff, Henrik
     11 APPLICANT: Liljestrom, Peter
     12 TITLE OF INVENTION: DNA Expression Systems Based on
     13 TITLE OF INVENTION: Alphaviruses
     14 NUMBER OF SEQUENCES: 27
     15
     16 CORRESPONDENCE ADDRESS:
     17 ADDRESSEE: Birch, Stewart, Kolasch & Birch
     18 STREET: P.O. Box 747
     19 CITY: Falls Church
     20 STATE: Virginia
     21
     22 COUNTRY: USA
     23 ZIP: 22040-0747
     24
     25 COMPUTER READABLE FORM:
     26 MEDIUM TYPE: Floppy disk
     27 COMPUTER: IBM PC compatible
     28 OPERATING SYSTEM: PC-DOS/MS-DOS
     29 SOFTWARE: Patentin Release #1.0, Version #1.25
     30
     31 CURRENT APPLICATION DATA:
     32 APPLICATION NUMBER: US/07/920.281C
     33 FILING DATE: 13-AUG-1992
     34 CLASSIFICATION: 435
     35
     36 ATTORNEY/AGENT INFORMATION:
     37 NAME: Murphy, Jr., Gerald M.
     38 REGISTRATION NUMBER: 28,977
     39 REFERENCE/DOCKET NUMBER: 828-103P
     40
     41 TELECOMMUNICATION INFORMATION:
     42 TELEPHONE: 703-241-1300
     43 TELEFAX: 703-241-2848
     44
     45 TELEX: 248345
     46
     47 INFORMATION FOR SEQ ID NO. 1:
     48
     49 SEQUENCE CHARACTERISTICS:
     50 LENGTH: 11517 base pairs
     51 TYPE: nucleic acid
     52 STRANDEDNESS: single
     53
     54 TOPOLOGY: linear
     55
     56 MOLECULE TYPE: RNA (genomic)
     57
     58 HYPOTHEetical: NO
     59
     60 ANTI-SENSE: NO
     61
     62 ORIGINAL SOURCE:
     63 ORGANISM: Semliki Forest Virus
     64
     65 FEATURE:
     66
     67 NAME/KEY: -
     68 LOCATION: 1..11517
     69 OTHER INFORMATION: /label= genome
     70
     71 OTHER INFORMATION: /note= "Semliki Forest Virus complete nucleotide
     72 sequence, presented as a cloned DNA sequence; see
     73 OTHER INFORMATION: Figure 5."
     74
     75 FEATURE:
     76
     77 NAME/KEY: CDS
     78 LOCATION: 87..7379
     79 OTHER INFORMATION: /product= "SFV polyprotein"
     80
     81 FEATURE:
     82
     83 NAME/KEY: CDS
     84 LOCATION: 7421..11179
     85 OTHER INFORMATION: /product= "SFV polyprotein"
     86
     87 US-07-920-281C-1
     88
     89 Query Match 31.3%; Score 1564; DB 1; Length 11517;
     90 Best Local Similarity 58.3%; Pred. No. 0;
     91 Matches 2904; Conservative 0; Mismatches 2040; Indels 36; Gaps 8
     92
     93
     94
     95
     96
     97
     98
     99
    100
    101
    102
    103
    104
    105
    106
    107
    108
    109
    110
    111
    112
    113
    114
    115
    116
    117
    118
    119
    120
    121
    122
    123
    124
    125
    126
    127
    128
    129
    130
    131
    132
    133
    134
    135
    136
    137
    138
    139
    140
    141
    142
    143
    144
    145
    146
    147
    148
    149
    150
    151
    152
    153
    154
    155
    156
    157
    158
    159
    160
    161
    162
    163
    164
    165
    166
    167
    168
    169
    170
    171
    172
    173
    174
    175
    176
    177
    178
    179
    180
    181
    182
    183
    184
    185
    186
    187
    188
    189
    190
    191
    192
    193
    194
    195
    196
    197
    198
    199
    200
    201
    202
    203
    204
    205
    206
    207
    208
    209
    210
    211
    212
    213
    214
    215
    216
    217
    218
    219
    220
    221
    222
    223
    224
    225
    226
    227
    228
    229
    230
    231
    232
    233
    234
    235
    236
    237
    238
    239
    240
    241
    242
    243
    244
    245
    246
    247
    248
    249
    250
    251
    252
    253
    254
    255
    256
    257
    258
    259
    260
    261
    262
    263
    264
    265
    266
    267
    268
    269
    270
    271
    272
    273
    274
    275
    276
    277
    278
    279
    280
    281
    282
    283
    284
    285
    286
    287
    288
    289
    290
    291
    292
    293
    294
    295
    296
    297
    298
    299
    300
    301
    302
    303
    304
    305
    306
    307
    308
    309
    310
    311
    312
    313
    314
    315
    316
    317
    318
    319
    320
    321
    322
    323
    324
    325
    326
    327
    328
    329
    330
    331
    332
    333
    334
    335
    336
    337
    338
    339
    340
    341
    342
    343
    344
    345
    346
    347
    348
    349
    350
    351
    352
    353
    354
    355
    356
    357
    358
    359
    360
    361
    362
    363
    364
    365
    366
    367
    368
    369
    370
    371
    372
    373
    374
    375
    376
    377
    378
    379
    380
    381
    382
    383
    384
    385
    386
    387
    388
    389
    390
    391
    392
    393
    394
    395
    396
    397
    398
    399
    400
    401
    402
    403
    404
    405
    406
    407
    408
    409
    410
    411
    412
    413
    414
    415
    416
    417
    418
    419
    420
    421
    422
    423
    424
    425
    426
    427
    428
    429
    430
    431
    432
    433
    434
    435
    436
    437
    438
    439
    440
    441
    442
    443
    444
    445
    446
    447
    448
    449
    450
    451
    452
    453
    454
    455
    456
    457
    458
    459
    460
    461
    462
    463
    464
    465
    466
    467
    468
    469
    470
    471
    472
    473
    474
    475
    476
    477
    478
    479
    480
    481
    482
    483
    484
    485
    486
    487
    488
    489
    490
    491
    492
    493
    494
    495
    496
    497
    498
    499
    500
    501
    502
    503
    504
    505
    506
    507
    508
    509
    510
    511
    512
    513
    514
    515
    516
    517
    518
    519
    520
    521
    522
    523
    524
    525
    526
    527
    528
    529
    530
    531
    532
    533
    534
    535
    536
    537
    538
    539
    540
    541
    542
    543
    544
    545
    546
    547
    548
    549
    550
    551
    552
    553
    554
    555
    556
    557
    558
    559
    560
    561
    562
    563
    564
    565
    566
    567
    568
    569
    570
    571
    572
    573
    574
    575
    576
    577
    578
    579
    580
    581
    582
    583
    584
    585
    586
    587
    588
    589
    590
    591
    592
    593
    594
    595
    596
    597
    598
    599
    600
    60
```

Dh 156 CCGTCGTTGAGTGGAGTCATGTCAGAGTCAACCAAAATGACATGCAATGCCAGACA 215
Qy 151 TTTTTCGATGTGGCAACAAAGCTCATTTGAGCGCAAGTCGACGGGACCAAGTTATCTTG 210
Db 216 TTTTTCGACCTGGGTACCAAAATGATCGAGAGAGACTGACAAAGACACATCATCTTG 275
Qy 211 GACATTTGAAAGTGGCCCGTGCAGATGCAATTCATTCACCGCTTCATCTTGTATCTGC 270
Db 276 GATATCGGAGTGGCGCTTCAGAGAAATGATGTCTACGCAAAATACCACTGGGTATGC 335
Qy 271 CTTATGATTAAGCGGTGAAGACCCGAGACACTTAACCGGTATGCAAGAAAGT-----T 324
Db 336 CCGTATGCGAGCGGAGAAAGACCCGAAAGGCTCGATAGCTACGCAAAAGAACTGGCAGCG 395
Qy 325 AAGAAAGTGCATTTACCGAAGAAACATTAACCTCTTAAGGGGAGACCTGCTGAAATGC 384
Db 396 GCGTCGCGGAAAGGTCTGGATGAGAGATCGCAAGAAATACCGACCTGCAACGCTC 455
Qy 385 ATGTCAACACGAGACGAGAGACTCATCTGTGTATGACACAGACGCGCATGTAG 444
Db 456 ATGGCTACGCGAGACGCTGAATCTCTTACCTTTGGCTGCAATACAGAGCTACGCTGTCT 515
Qy 445 TACTTTGAAAGTGTAGCAGTATACAGAGATGTATACGAGTCCATGCAACGACATCAATC 504
Db 516 ACGGCAACCGAAAGTGGCGGTATACAGAGACGTGTATGCTGTATACATGCAACAACTCGCTG 575
Qy 505 TACCAACGAGGCTTTAAAGAGTTAGAGCAATTTTACTGATAGGCTTTGACACGACCCCT 564
Db 576 TACCATAGGCGGATGAAGAGGTGTAGAAACGCGGTATGGATTTGAGTTGACACACCCG 635
Qy 565 TTTATGTACAAAAACATGGCAGGTCTTCTTCACTTCAACAACGAAAGTGGGCTGACGAG 624
Db 636 TTTATGTGTGACGCGCTTACGAGCGCGCTTATCCAACTTACGCAAAACTGGGCGAGAG 695
Qy 625 AGAGTATTTGAAAGCAGTAACTTGGCGCTCGTAACTCAATCTTCAGAGAGCAGGCTT 684
Db 696 CAGGTGTACAGGCAAGAAACATAGACTGTGTGACAGATCTTGAATGAGGAAAGATC 755
Qy 685 GGAATATCTCAATCTTGAAGAAAGAGGCTCCAACTTAATTAAGTATATTTCTG 744
Db 756 GCGAACTGTCAATCTTCCGCAAGAAAGATTTGAACCTTGCACACAGTATTTCTG 815
Qy 745 GTTGTTCACAAATCTACACAGAGATGATCACTGTTACGTAGCTGGCATCTTCCAAC 804
Db 816 GTAGGATCTACATTTGACACTGAGAGCAAGAAAGTACTGAGGAGCTGCACTTACCTCC 875
Qy 805 GTGTTCACCTTGAAGAAAGTCTTAATCTCAAGGTAGATGTGGGACCATTTGACGCTGT 864
Db 876 GTATTTCACTGAAAGGTAAACATCTTACTGTAGGTGCAATCCATGATGATCATGT 935
Qy 865 GAAGGATACGTCACTCAAAAGATTAACGATCAGCCGAGACTTATACGTTAAAGTTGAGAC 924
Db 936 GAAGGATACGTATTAAGAAATCACTATGTGCCCGCTGTACGGTAAACGGTAAAGG 995
Qy 925 TTGGCGTCAACAATGCAATCGGAGGAGTTTCTTGAATTGCAAAAGTCAAGATACGCTGGC 984
Db 996 TAGCGCTGAGCTATCAAGGAGGAGGATTTCTAGTGTCAAGACACAGACACTGTCAA 1055
Qy 985 GCGGAGAGGTTTCTTTGTGTGTGTATCGTATGACAGCCACACTTTGGCATCAATG 1044
Db 1056 GGAAGAAAGTCTCATCTTCCGTATGCACTTACGCTTCAACCATCTGTGATCAATG 1115
Qy 1045 ACAGGAGTCTGGCACTGAGAGTTAGTGTGATGACGACAAAACTATTTGGTGGGCTC 1104
Db 1116 ACTGGCACTAGGACCGAGCTCAACCGAGAGACGACAGAAAGTTGTGTGGGATG 1175
Qy 1105 AACCAAGATTTGCTCAATGTAGAGCGCAAGAAATACTATCACAATGACAGACTAT 1164
Db 1176 AATGAGAGATAGTTGTGAACGAAAGACACAGGAAACCTTAACAGATTAAGACTAT 1235
Qy 1165 CTATTACAGTGTGTGCGCCAGGCGTTTTCAGGTGGGCGCTGAACTGTGTCGCACTTG 1224
Db 1236 CTGCTTCGATTTGTGGCGGTGGCATTTTACAAAGTGGGAGGAAATCAAGCAGACCTT 1295

Qy 1225 GAGAGAGAAAGAACTTAAAGGGGTGGGAGGCACTTTACTATAGGCTGTGCTGGCT 1284
Db 1296 GATGATAAAAACCTCTGGGTGTCCGAGAGGTCACTTACTTGTCTGTGTGGGCA 1355
Qy 1285 TTCAAGACCCAGAAATACATCCATCTTACAGAAAGCTGTGACGCAAACTTAAAGAA 1344
Db 1356 TTTTAAACGAGAAAGATGCAACCATGTACAGAAACCAAGACCAAGCAATGTGAG 1415
Qy 1345 GTACTGCGCTCTTTGACTCATTTGTGATTCACGCTTACAGCCAGCGGCTGATATG 1404
Db 1416 GTGCTTTCAGAGTTTAACTGTGTGTATCCGAGCTTATGTCTTACAGGCTTGCATC 1475
Qy 1405 GCGTTCGCGGTAGGCTCAAGCTGCTTGAACCAACTGCAACCCGACCGGCTATT 1464
Db 1476 CCACTCAGATCAGCATTTAAAGATCTTTTGGCCAAAGAACCAAGC--GAGATTATA 1532
Qy 1465 ACAATGCGCATGTGAGCATCTGCGGTTCACAGCAAGAGCTGAAGAGTGTGCA 1524
Db 1533 CTTGTTCTGACCGGTGTGACGACGAGGATGTGAAACAGAGGAGAGAGTGTGAG 1592
Qy 1525 GCGGAAAGATCAGAGAGCCCTGCAACCTTGTCTCCTGAATGAAAAAGAGCCGTA 1584
Db 1593 GCGGAGCTGACTAGAGAGGCTTACACACCTCGTCCCAATCGCGCGGAGAGACGGA 1652
Qy 1585 GAGGAGAGTAAAGCTTATGCAAGAG-----GACAGAGCAGGTAGGTGAGACA 1638
Db 1653 GTCTGCAAGCTCGAGCTTTGAAGAACTTAGATATCACAGAGTGCAGGGGTGTGGAACA 1712
Qy 1639 CCAAGAGACACATCAGGAGTGAACAAAGTTACCCAGGCGAAGAGATTTGGTCTTACGCT 1698
Db 1713 CTTGCGAGCGGTTGAAGATCACCGCACAGCCGAAAGCATACTAGGAAATTAAGTA 1772
Qy 1699 ATACTTCACTCCAGGCGGTATGTGAATGTGAAAACTGCGGTGTATCCACCACTTGGCG 1758
Db 1773 GTTCTGTCCCGGAGAGCGGTGTCAAGAGCTCCAAAGTTGCCCCCGGTGCAACCTCTGCA 1832
Qy 1759 GAACAACTACTGGAATGATCTCAAAAGGTAGGGCAGGAGATTAACAAGTGCAGCCATC 1818
Db 1833 GAGCAGGTGAATTAATTAACAATTAACGAGAGGCGCGGTTTACAGGTGACGAGTAT 1882
Qy 1819 CACGTAAGTCAATTTGTACAGAAAGGACGCGGTCCCTGTTCAGAACTTCCAGCATTG 1878
Db 1893 GAGCGAGGCTCTACTACATGATGATGATCGGCATTCGGGTCCGTGATTCAAGCTTTG 1952
Qy 1879 AGTGAAGCGCTACGATCTTTTCAACGAGAGGAGTTGTTAAACAGATACCTGCACAC 1938
Db 1953 AGCGAGGCGCACTATGTGTATCAACGAAAGGAGTTGTCTCAACAGAAACTATATCAT 2012
Qy 1939 ATGCAATCAACGAGAGGAGGCTTAAACATGACGAGAGATCTATTAAGCTGTAAAGCT 1998
Db 2013 ATTTGCTTACAGGACCGTGTGTBAACCGAGAGGAACTTACGAAAGTCAAGCT 2072
Qy 1999 CAGGACACAGACTCAGAAATAGTCTTGTGATTTGACGCAAGAAAGTGTGTAAGCGAGAA 2058
Db 2073 GAAAGAACTAAGCGCGAGTACGTGTGACGTGATTAAGAAAGTGTGCTGCAAGAGAG 2132
Qy 2059 GAGCGAGGCTCTTGTGCTTAAACCGGTGTGTGTATGATCAACATTTCAAGTTTGGC 2118
Db 2133 GAAGCGTGGGTTTGGTGTGTGAGAGACTTAACCAACCCCGTGTCAATGATTTGCC 2192
Qy 2119 TAGGAGTCTCAAGACAGACGACGACCTTCAAAAGTCCCAACCATCGAGTAT 2178
Db 2193 TACGAGGCTGAAGATCAGGCGGTGCGCACATTAAGACTTAAGTATGAGGTCTTT 2252
Qy 2179 GAGTGCAGGTTCAAGTAAATCTGATCATCAAAAGCGCTGTGACTAAGAAAGATCTG 2238
Db 2253 GGGGTTCCGGGATACGCAAGTCTGTATTAATTAAGGCTGTGACCAACAGATCTG 2312
Qy 2239 GTTGTAGTCCGAGAAAGAAACTGCGCAGAAATCATCAGGATGTAAAGAGATGAGA 2298
Db 2313 GTCAACGCGGCAAGAGGAACTGCGCAAGAAATAGTTAAAGCATGGAAGAAACCGC 2372

QY 2299 CGATGATGTTGCTGCTAGAGCTGTGCAATTCAGTCTTTAAATGAGGTTAAGCACCC 2358
Db 2373 GGGAAAGGGGACAAAGTAGGAAAAAGTAGCTCATCTGCTTAAACGGGGTGTGCTGTGCC 2432
QY 2359 GTTAACTCTGTACATTGATGATGAGCATTTGCTTGCATGACAGGACGCTGTGCACCTG 2418
Db 2433 GTGACATCTTAATATGAGACAGGCTTTCCTTCCCATTCGCGTAATCTGTGTCCTTA 2492
QY 2419 ATTGCGATCTGCTAAACCT---AAGAAAGGTATGTGTGGGGAGCCCAAAACAATGGCGC 2475
Db 2493 ATTGCTCTTTTAAACCTCGAGACAAAGTGTGTATGTGGAGACCCCAACAATGCGGA 2552
QY 2476 TTTCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2535
Db 2553 TTTCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2612
QY 2536 CATAAAGCATCTCTAGAGAGGTGACACAGATGTAAACCGCATGTCTTCCACGCTTTC 2595
Db 2613 CATAAAGCATCTCTAGAGAGGTGACACAGATGTAAACCGCATGTCTTCCACGCTTTC 2672
QY 2596 TACGACAAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2655
Db 2673 TACGACAAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2732
QY 2656 GGGACCAAAAGCCGACAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2715
Db 2733 GGGACCAAAAGCCGACAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2792
QY 2716 CAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2775
Db 2793 CAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2852
QY 2776 ACGGGAAGGGGTTATGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2835
Db 2853 ACGGGAAGGGGTTATGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2912
QY 2836 ACTTGTGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2895
Db 2913 GCGTGTGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2972
QY 2886 CTAGTGTGATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2955
Db 2973 CTGCGCGGATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3032
QY 2956 TCATTGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3015
Db 3033 ACATTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3092
QY 3016 CAGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3075
Db 3093 GCGCTGTGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3152
QY 3076 GTCTTGTGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3132
Db 3153 GTCTTGTGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3212
QY 3133 TTTCAAGCATGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3192
Db 3213 TTTAAAGGAGCAGAGCTTACTCTCAAGTGTGCTTGAATGAAATTTGACCAAGTAC 3272
QY 3193 TTTGAGATGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3252
Db 3273 TATGAGATGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3332
QY 3253 GATGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3312
Db 3333 AACAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3392
QY 3313 AAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3372
Db 3393 GCGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3452
QY 3373 GATATGAAGATTAATACATCAAGACTACTCTCAACAATTAATGATGATGATGATGAT 3432

Db 3453 GTATTCGACAAAGAAAAATCAACCGCTTCTGTGCTGACAAATGATTAATCTTCAAC 3512
QY 3433 CGCGGTTGCCCCACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3492
Db 3513 CGCAGGCTGCGCACCGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 3572
QY 3493 GGAATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3552
Db 3573 TGGCTGTGATTAAGTAAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 3632
QY 3553 CGAGGAAGAAAGTAAAGTAAAGGATGATGATGATGATGATGATGATGATGATGATGAT 3603
Db 3633 TTTGCTGACAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3692
QY 3604 TGTATCTGATTTGGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3663
Db 3693 TACGACCTTAAGTAAAGTAAAGGATGATGATGATGATGATGATGATGATGATGATGAT 3752
QY 3664 GTTAGGACCCGTTACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3723
Db 3753 ATTACACAGGATTAAGTAAAGGATGATGATGATGATGATGATGATGATGATGATGAT 3812
QY 3724 AGCATGCTTAAGTAAAGTAAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 3783
Db 3813 CAGATGCTTGGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3869
QY 3784 GGGATGAGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3843
Db 3870 GCTTACGATTAAGTAAAGTAAAGGATGATGATGATGATGATGATGATGATGATGATGAT 3929
QY 3844 AGGTTTACCGGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3903
Db 3930 TGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3989
QY 3904 TTTCTGCGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3963
Db 3990 TTTCTCAACTTTGACAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4049
QY 3964 GACAACTATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4023
Db 4050 AGTCCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4109
QY 4024 ATCAGAGTGAATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4083
Db 4110 AAGAGAGACATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4169
QY 4084 GGTCAACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4143
Db 4170 GGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4229
QY 4144 AGACAGCAATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4200
Db 4230 GGAACAGCAATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4289
QY 4201 CATGCTTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4260
Db 4290 CACGCTTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4349
QY 4261 GCTGCTTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4320
Db 4350 GCTGCTTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4409
QY 4321 CGGCTTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4380
Db 4410 CGGCTTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4469
QY 4381 CACGCTTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4440
Db 4470 CATCTTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4529
QY 4441 CAATGAGACCAAGATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 4500

4530 AGTTGGAGAGAAAATTCAGAGAACATTCATGAGACGCGTGTGAGTTGCTCAAT 4589
4501 GATGACAGACGATGACATTCATGCTGAGGGTCCACCAACAGCTCTTTGGCAGGC 4560
4590 GATGACGTGAGCTGACACGACCTGCTGAGAGTGCACCGGACAGACGCTGCTGGT 4649
4561 AGACAGGTTACTCCGTCATAGAGGCAAGTTATTCATACCTGGAAGTACAGATTC 4620
4650 CGTAAGGCTTACATACCTGACGAGCGGCTGTACTGTCTTGAAGGTACCAATTC 4709
4621 CATCAGACCGCCAGACATTCGCGAAATCCATGATGAGGCCCAAAATCTGAGGCT 4660
4710 AACGAGCTGCTATGATGATGAGAGATCTGACGTTGTGGCCACAGCTCAGAGGCA 4769
4681 AATGAGAGATTTCTGTTATCATCTGCGGAGATATGTCACATCTCCGTTCCAAATGC 4740
4770 AACGACAGATATCTTATGACGCGCTGGCCAAACAAATGACAACTGATCAAAATGT 4829
4741 CCAATGAGAGATGAGAGCGCTGCTCCTCACCTCACACTTCATGCTGTAAATTAC 4800
4830 CCGGTGAACATTCGATTCATCAACACTCCAGACAGTCCCTGCTGTGCGCTAC 4889
4801 GCTATGACGCTGACCGCGTATACAGTTGCTGCTGCGAAAGAAACAGTTGCGCTA 4860
4890 GCAATGACAGACAGACGATCGCCGCTTAGTTCACACCAAGTTAAAGCATGCTGTT 4949
4861 TGTCTCATCATCTCTGTTGCGCAATACAGATTCACAGCGCTGCGAAGCTACATGACG 4920
4950 TGTCTCATCT 5009
4921 AAACAGCT 4980
5010 AAGGTTCT 5069

RESULT 5

US-08-466-277-1

Sequence 1, Application US/08466277

Patent No. 6190666

GENERAL INFORMATION:

APPLICANT: Garoff, Henrik

Liljestrom, Peter

TITLE OF INVENTION: DNA Expression Systems Based on

Alphaviruses

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESSES:

ADDRESS: Birch, Stewart, Kolasch & Birch

STREET: P.O. Box 747

CITY: Falls Church

STATE: Virginia

COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,277

FILING DATE: 06-Jun-1995

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/920,281

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Murphy Jr., Gerald M.

REGISTRATION NUMBER: 28,977

REFERENCE/DOCKET NUMBER: 828-103P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-241-1300

TELEFAX: 703-241-2848

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 11517 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Semliki Forest Virus
FEATURE:
NAME/KEY: 1..11517
LOCATION: 1..11517
OTHER INFORMATION: /label= genome
/note= "Semliki Forest Virus complete nucleotide
sequence, presented as a cloned DNA sequence; see
Figure 5."

FEATURE:
NAME/KEY: CDS
LOCATION: 87..7379
OTHER INFORMATION: /product= "SFV polypeptide"

FEATURE:
NAME/KEY: CDS
LOCATION: 7421..11179
OTHER INFORMATION: /product= "SFV polypeptide"

SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-466-277-1

Query Match 31.3%; Score 1564; DB 3; Length 11517;
Best Local Similarity 58.3%; Pred. No. 0;
Matches 2904; Conservative 0; Mismatches 2040; Indels 36; Gaps 8;

31 AGATTACCTGATCTTAAAGTGTGACAGCGCCGATGCTCAAGTGTGACAGCGGACGTT 90
96 AAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 155
91 CCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 150
156 CCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 215
151 TTTTGCATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 210
216 TTTTGCATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 275
211 GACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 270
276 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 335
271 CTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 324
336 CTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 395
325 AAGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 384
396 GCTCCGGAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 455
385 ATGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 444
456 ATGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 515
445 TACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 504
516 ACGGACCGAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 575
505 TACCAAGGCGCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 564
576 TACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 635
565 TTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 624
636 TTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 695
625 AGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 684

Db 696 CAGGTTACAGCGCAGGAAACATAGACGTGTGACGACATCCCTTGAAGGAAAGACTC 755
Qy 685 GGAATCTCTCAATCTCTTAGAAGAAAGAGCTCCAACTCAATTAATTAATCTATTTCTG 744
Db 756 GCGAATCTGTCTCTCCGCAAGAAAGCAATTGAAACCTTGACACAGTCAATGTTCTCG 815
Qy 745 GTTGTTCACAAATCTACACAGAAATAGATCACTGTTTACGTAGCTGACATCTTCCAAC 804
Db 816 GTAGATCTACATGTAGACCTAGAGCAAGAAAGCTACTAGAGAGCTGGACCTTACCTCC 875
Qy 805 GTGTTCCACTTGAAGAAAGTCTTAATTCAAGGTAGATGTGGGACCATTTGTCACTGT 864
Db 876 GTATTCACCTGAAAGGTAAACAAATCTTTACCTGAGGTGCAATACCATCTATCATGT 935
Qy 865 GAAGGTCGTGATCAAAAAGATTAAGATCAAGCCAGGACTATAGGTTAAATTTAGAAC 924
Db 936 GAAGGTCGTGATTAAGAAATCACTATGTCCCCCGCTGTAGCGTAAACGGTAAAG 995
Qy 925 TTGCGTCACAAATCATCGCAGAGGTTTCTTGAATGCAAAATCAAGATACGCTGCG 984
Db 996 TACGCGTAGACATCAACGGGAGGATTCCTAGTGTGAAACACACAGACACTGTCAA 1055
Qy 985 GCGAGAGGTTTCTTTGCTGTGTACGATGTACAGCCACCACTTTGCGATCAGATG 1044
Db 1056 GGAGAAAGTCTCATTCCTGTATGACCTACGTCCCTCAACCATCTGTATCAATG 1115
Qy 1045 ACAGGATTTCTGCACTGACCTTGTGTGTAGTGTACAGCAAAATCTATGTTGGGCTC 1104
Db 1116 ACTGCATCTAGCAACGACCTCAACCGAGAGACGACAGAAATGTTTATGTGGATG 1175
Qy 1105 AACCAAGGATTTGTCATAGTGTAGAGACGCAAAAGAAATCTATACAAATGACAACTAT 1164
Db 1176 AATCAGAGATAGTGTGACAGGAAAGAACACAGCAACCTAACATGTAAGAACTAT 1235
Qy 1165 CTATTAACAGTGTGTGCGCCAGGCGTTTTCAGGTGGCGCGTGAACATGTCGCACTTG 1224
Db 1236 CTGCTCCGATTTGTGGCGTGCATTTAGCAAGTGGGAGGAAATCAAGGACACCTT 1295
Qy 1225 GACGCGAAGAAAGACTAGGGGTGGGAGAGCCACTTTACTATAGGGCTGTCTGGGCT 1284
Db 1286 GATGTGAAAAACCTCTGGGTGTCCGAGAGAGTCACTTCTGTGTCTGTGGGCA 1355
Qy 1285 TTCAAGACCCAGAAATCACTCATCTACAGAAAGCTGTGACCAACAAATTAAGAA 1344
Db 1356 TTTAAAGAGAGAAATGACACCATGTACAAAGAAACAGACACCAACAAATTAAGTAA 1415
Qy 1345 GTACCTGCGTCTTGTGACTATTTGTGATTCACGCTTACAGGACGCGCTCGATATG 1404
Db 1416 GTGCTTCAGAGTTTAACTCGTTCGTCAATCCGACCTATGTGTACAGGCTCGCAATC 1475
Qy 1405 GGCCTTCCGCGTAGGCTCAAGCTGTGTGTTGAACCACTGTAAACCCGACCGCTATT 1464
Db 1476 CCAAGCAATCAACGATTAAGATGTCTTTTGGCCAAAGAACCAAC---GAGAGTTAATA 1532
Qy 1465 ACAATGCGCATGTGAGCATCTGCGTGGCTTACAGCAAGAGCTGAAGAGTGGCTCA 1524
Db 1533 CCTGTTCTGACGCTCGTCAAGCCAGGATGTGAAACAAAGAGAGAGAGGTTGAAG 1592
Qy 1525 GCGGAAGATCAGAAAGCCCTGACCTTGTCTCCTGAAATGAAGAAAGAGCCGTA 1584
Db 1593 GCCGAGCTGACTAGAGAGCTTACACCCCTCGTCCCATGCGCCGCGGAGAGCGGA 1652
Qy 1585 GAGCGAAGATGAGCTCATTAATGCAAG-----GACAGACAGGTAGGTGAGACA 1638
Db 1653 GTCTGCGAGTGAAGTGAAGAACTTAGATGTACGACAGGTGCGGGTCTGTGAAACA 1712
Qy 1639 CCACGAGACATCAGGATGACAGTTAACCGAGCGAAGAGAAATGTTGGTCTTACGCT 1698
Db 1713 CCTGCGACGGGTTGAAGTCAACCGCACAGCGAAAGAGCTACTAGGAAATTTACGTA 1772
Qy 1699 ATACTTCAACCCGAGCGGTATTTGAATAGTGAAGAAATGCGCGTGTATCAACCATTTGCG 1758

Db 1773 GTTCTGCCCGCAGACCGGTCTCAAGAGCTTCAAGTTGGCCCCCGTGACCCCTTAGCA 1832
Qy 1759 GAACAAGTACTGTAAATGATCTCAACAAGTATGGGACGGGAGATAGAAATGAGCATAC 1818
Db 1833 GAGCAGGTGAATATATACATTAACGAGAGGCGCGGCTTACAGGTCAAGCAATAT 1892
Qy 1819 CACGTAAGTCTATTTGATCCAGAAAGGAGCGCGGCTCCCTGTTCAAGACTTCCAGCATTTG 1878
Db 1893 GACGCGAGGTCTCTATCAATGTGTGATCGGCATTTCCGGTCCCTGAGTTCAAGCTTTG 1952
Qy 1879 AGTAGAGCGCTACATGTTTCAAGAGAGGAGTTGTGAACAGATACCTGCACAC 1938
Db 1953 AGCGAGACGCGCATATGTGTGAACAAAGAGGAGTTGTGTCACAGAAACTATATACAT 2012
Qy 1939 ATCCGAATCAACGAGAGCGGTAAACCTAGACGAAAGTACTATTAAGCTTAAGACT 1998
Db 2013 ATGCTGTTCAACGAGCGCTCGTGAACACGAGGAGAACTTACAGAAATCAGAGCT 2072
Qy 1999 CAGACACAGACTCAGAAATAGCTCTTCAATATTTGACGACGAAAGTGTGTTAAGCAGAA 2058
Db 2073 GAAGAATGACGCGAGTACTGTGTTCAAGTAAATTAATGCTGCTCAAGAGAGAG 2132
Qy 2059 GACCGAGTCCCTGTGTCTTAAACCGGTATCTGTAGATCCAACTTTCAGGATTTGCG 2118
Db 2133 GAAGCGTGGGTTGTGTGTGTGGAGAGCTTAACCAACCCCGTTCATGATTTGCGC 2192
Qy 2119 TACGAGATCTCAAGACAGACAGACAGACCTTCAAAAGTCCCAACCATGGAATCTAT 2178
Db 2193 TACGAAGGCTGAAGATCAGGCGCGTCCACATTAAGACTACAGTATGAGATCTTT 2252
Qy 2179 GAGAGTCCAGGTTCAAGTAAATCTGGAATTCATCAAAAGCGCTGTGACTAAGAAAGCTG 2238
Db 2253 GGGGTTCCGGATCAGGCAATGTCTTATTTAAGACCTCTGTACCAACCAAGATCTG 2312
Qy 2229 GTTGTAGTCCGAAGAGAAACCTGCGCAGAAATCATCAGGATGTAAAGAGATGAGA 2298
Db 2313 GTCAACAGCGGCAAGAAAGAACTGCGCAGAAATATGTTAAGAGTGAAGAAAGCACGCG 2372
Qy 2299 CGTATGATGTGCTGTAGAGACTGTGATTCAGTGTCTTCTTAAATGGGTTAAGACCC 2358
Db 2373 GGGAAAGGAGACAGTGAAGGAAACAGTACATCCCTGTAAACGGGTGTGTGTGCTC 2432
Qy 2359 GTTAACTCTGTATGATGATGAGGATTTGCTGCTGCAATGAGGAGCGCTGCGACATG 2418
Db 2433 GTGACATCTCTATATGTGGAAGAGCTTTCGCTTCCATTCGATCTCTGTGCGCTTA 2492
Qy 2419 ATTGCATGTCAACT--AAGAAAGTGTATTTGTGCGGAGACCCAAACAAATGCGGC 2475
Db 2493 ATTGCTTTGTTAAACCTCGAGCAAGTGTGTATGCGAGACCCCAAGCAATGCGGA 2552
Qy 2476 TTTCTTAACTGATGTGCTGAAAGTACATTTTAACTGACATATGCACTGAAATGTAC 2535
Db 2553 TTTCTCATATGATGACAGCTTAAGGTGAATTTCAACCAACCAATCTGCACTGAAGTATGT 2612
Qy 2556 CATAAAGCATCTCTAGAGGTGACACAGACTGAACCGGCATGCTCCACGCTCTTC 2595
Db 2613 CATAAAGTATATCAAGCTTGAACCGCTCAAGTCAAGCGCATGTCTTACCTTGCAC 2672
Qy 2596 TACGACAGCGAATGAAGACGTTAACCCATGTGCTGTAATAATCATATGATATACCA 2655
Db 2673 TACGAGGACAGATGCGGACAGCAACCCGTCACAAACCAATATATATGACACCA 2732
Qy 2656 GGGACACAAAGCCGACAAAGATGATTTTAACTTCTGTTTCAAGAGATGGTGA 2715
Db 2733 GGAACAGACCAAGCCCAAGCCAGAGACATCTGTATTAATCTTCCGAGGCTGGCAAG 2792
Qy 2716 CAGCTACAGATTGACTCAAAATCAAGAAATCAATGATGCGGCTGACGCAAGGACTT 2775
Db 2793 CAGCTGAGTTGACTACCTGTGACAGAAATCAATGACAGACAGATCTCAAGGCTTC 2852
Qy 2776 ACAGGAAAGGCGTTTATGTGTGAGTCAAGATCAACGAAATCACTACTGCGAG 2835
Db 2853 ACCGCAAGGGGATATAGCGCGTAAAGGCAAGAGTGAATGAATTCCTGTATGCGCCT 2912

OY	2836	ACCTTCAGGACACGTTGAACGTGTCTCTTACAGCACAAGAAACGATGTCTGGAAGCG	2893
Db	2913	GGGTGGGACACGTGAATGTACTGTGTGACGGCATTGGATGAGCTGTGTGGAAAACG	2972
OY	2896	CTAGCTGTGTATCCCTGGATTAAGACATTACAGCTAAATATATCCCGGGGATTTTCA	2955
Db	2973	CTGGCCGGCGATCCCTGTGATTAAAGTCTCTATCAAACTTCCACAGGGTATCTTTAC	3032
OY	2956	TCATTGGACGACTGGCAGCGGGAACAACAAGCCATTATGGCAGCGTCTTGATTAAGCG	3015
Db	3033	ACATTGGAAAGATGGCAGAAAGAACACACAATAATGAAGTGATTAAGAACCGGCT	3092
OY	3016	CAGACAGCTGATGTGTCTCAGAAATTAAGGTGAACGCTCTGGGCGAAGGCTTTAGAGCA	3075
Db	3093	GGCCTGTGGACGCGTTTCCAAACAAAGCGAATGTGTGTGGCGAAAAGCTGTGCTCT	3152
OY	3076	GTCCTTGGCCACGGCCCAACATTGTGTGTGACGAGACAGACAGTGGAGAC---GTTCAC	3132
Db	3153	GTCCTGGACACTGCGCGGAATCAGATTGACACAGAGGAGTGGAGCACCATTAATTA	3212
OY	3133	TTCAAGCATGACAGAGGCTACTCACCTGTAATGGACACTGAATCTCTTTGGACACAG	3192
Db	3213	TTTAAGGAGACAGAGCTTACTCTCAATGTGTGGCTTGAATGAATTTGACCAAGTAC	3272
OY	3193	TTTGGAGTAGACCTGGACAGTGGGTTATTTTCCGCTCTTACCGTTCGACCTTACTTA	3252
Db	3273	TATGAGATTGAACCTGGACAGTGGCTGTGTTTCTGGCCCGAAGGTGTCTCTGATTA	3332
OY	3253	GATCAGCATCTGGGATTAATCTCCGCGAAGGAACAATGTATGGGCTTAATAGAGAGT	3312
Db	3333	AACAAACACTGGGATTAACAGACTGTGTGGAAGATGTGATTCAAATCCGCAACACT	3392
OY	3313	AAGAGTTGTACACGGCGATATCCGTGCATCACAAAAGGGTTGACACAGCAGGGT	3372
Db	3393	GCCAGGCTTGAAGCTAAGCATATCTTCTGTAAGGGCAGTGGCATACGGCGAAGAGCA	3452
OY	3373	GATATTAAGGAATATACATCATCAAGACTACTCTCAACAATTAATGTGGTTCCATTAA	3432
Db	3453	GTTATCCGAGAAAGAAAATCCAACCGCTTTCTGTGCTGACAAATGTAAATCTATCAAC	3512
OY	3433	CGCGGGTTGCCCACTCGTTGATCGTTTACCAAAAGGACAGGGGTACAACATGATCACAGC	3492
Db	3513	CGCAGGCTGCGCACCGCTGTGTGCTAGTACACAGCGTTAAAGCAGTAGGGTTGAG	3572
OY	3493	GGATTCTATCTAAGATGAAGGGAATCTGTGTGTGATCGCGGATCTATCAGACTT	3552
Db	3573	TGGCTGTCAATTAAGATTAAGAGGGTACACAGCTCTGTGTGAGTGAATCAACTGGCT	3632
OY	3553	CCAGGGAAGAAATGATGAGTCCATGGGTCCATTGGCCACTAATAC-----CATCAGG	3603
Db	3633	TTGCTTCAGCGCAGGGTCACTTGTGTGTACCGCTGAATGTACAGGGCCGATAGGTGC	3692
OY	3604	TGTGATCTGCATTTTGGGAATACCTGACCATGTCCGGTAAATATGACATTATCTTTGTCAT	3663
Db	3683	TACAGACTTAAGTTAGACTGCGCGCTGACCGCGCAGTTTCGACTTGTGTGTGAC	3752
OY	3664	GTTAAGACCCCGTACAGAACCATCACTACCAACGTGCGAGGATCACGCTATCCACAC	3723
Db	3753	ATTCCACAGCAATTCAGATCCACACATACACAGCAGTGTGTGCACACGCGCATGAAGCTG	3812
OY	3724	AGCANTGCTAACGTGTGAAGGCTGTCCACACCTGGAACCTGGCGGAACATGTGTGCTATA	3783
Db	3813	CAGATGCTTTGGGGGAGATGCGCTTACATGCTTAAACCCGGGGGCAATCT---TGTGAGA	3863
OY	3784	GGGATATGGGCTTGTGATTCGGCAACCGAGAAATATCATCACTGCGGTGGCAGGCTCATTT	3843
Db	3870	GCTTACGGAATACGCGGATTAATAATCAGCGAAAGCGTTGTTTCTCTTAAAGCAAGAAATTC	3929
OY	3844	AGGTTTACCCTGTGTGTTCAGCTTAAGAACACTGCCAAAATATCTAGAGGTTCTCTTGGTG	3903
Db	3930	TGCTGTGCAAGAGTGTTCGCCCGGATTTGTGTCAACACCAATTAACGAAGTGTCTTGTGTG	3989

QY	3904	TTCTTTGGCAAGGCAACCGCAACCAACAACATCAACAGACACAACTCGGTGTAGTCTT	3963
Db	3990	TTCTTCCAACTTTGCAACGAGGAAAGAGACCTCTACGCTACACAGATGAATACCAAGCTG	4049
QY	3964	GACAAACATCTATCAAGGGTCTCAACCAAGATACAGAGCAGAGAGAGCTCCAGCTACAGATG	4023
Db	4050	AGTGCCTGTATGCGGAGGAAGCCATGTGACACGGCCGGGTGTGTCAACCATCTTACAGATT	4109
QY	4024	ATCAGAGGTGACATTAGCAAGAGCGCTGACCAAGCTATCGTTAATGCTGCTTAATAGCAA	4083
Db	4110	AAGAGAGCAGACATAGCCACGTGTGACAGAAAGCGGCTGTGTAAACGACGTAAAGCCCGT	4169
QY	4084	GGTCAACAGAGTTCCGGAGTGTGGCGGTGCACTGTACCGAAAAATGGCCGGCTCTTGTAT	4143
Db	4170	GGAACTGTAGGGAGATGGCGTATGTGACAGGGCCGTGGCGAAATATGGCTGTACCCGTATC	4229
QY	4144	AGACAGCCAAATAGGTGTCTGGAGCGGCTAGACTTGTGAAGCAGAACCG--CTCATATA	4200
Db	4230	GGAGAGGCAACACAGTGTGGGCACAATTAAACATGATGTGCGGTGTACCCCGTATC	4289
QY	4201	CATGCTGTAGAACCCCAATTTTTTTCTTAAGATGCCGGAACCGAGGGCGACCTTAAGCTGCA	4260
Db	4290	CACGCTGTAGCGCTTAATTTCTCTGCAACGACTAAGCGGAAGGGGACCCGGAATTGGCC	4349
QY	4261	GCTGCTCATAGAGATGACGTCCATCGTCAACCGCTGACGGGATTAACAAAATATCAGTA	4320
Db	4350	GCTGTCTACGGGCGAGTGGCCGGGAAGTAAACAGACTGTATCTAGCAGCGTGAACATC	4409
QY	4321	CCGCTACTGTCAACCGGATCTATTTCTGTGTGGCAAGATCGATGATCAATCATTTGCAT	4380
Db	4410	CCGCTGCTGTCCAAGAGTGTTCAGCGCGCGGAAGAGATAGCTGTGACGAATCCCTCAAC	4469
QY	4381	CACCTGTTCATCTGCTTTGGAACCTAAGGATGCGGCATGTACATATATGTGTTGATTA	4440
Db	4470	CATCTATTCAAGCAATGTGACGCCAGCGACGCTACGTGACATCTACTGAGAGACAA	4529
QY	4441	CAATGGGAGACAGAGATATGAGAGGCCATTACCGCAAAAGAAACGTCGAATTTCTGGAT	4500
Db	4530	AGTTGGAGGAAGAAATTCAGGAAGCCATGTACATGAGAACGGCTGTGTGATGTCTCAAT	4589
QY	4501	GATGACAAAGCAGTAGACATTGACTTGGTCCAGGCTCCACCCAAACGCTCTTTGGCAGGC	4560
Db	4590	GATGACGTGTGAGCTGACCAACAGACTTGGTGAAGATGTGACCCGGACAGAGCCTGTGGGT	4649
QY	4561	AGACCAAGTTACTCCGTAATGAGGGGAAGTTGATTTCAATCTCGGAAGGTACAGATTC	4620
Db	4650	CGTAAGGCTACAGTACCTGACGAGGGGTGCTGTACTGTACTTTGAAGGTACCAATTC	4709
QY	4621	CATCAGACCGGCAAGACATTGCCGAATTCATGCAATGTGGCCCAACAAATCTGAGGCT	4680
Db	4710	AACCAAGGTGTATTTGATATGGCAAGATATCTAGCGTTGTGGCCACAGCTGCAAGAGCA	4769
QY	4681	AATGAGCAGATTTGCTTTATACATCTGGGGAGATATGTCACAGATTCGCTCAAAATGC	4740
Db	4770	AACGAAACGATATGCTATATACGCCCTGGGCAAAACATGTGCAACATCTAGATCCAAATGT	4829
QY	4741	CCAGTAGAGAGTCAAGAGCGCTGTCTCACTCAACACTTCATGCTGTGTAAATAC	4800
Db	4830	CCGGGAACGATTCGATTCAATCAACCTCCAGAGAAAGTGCCTGTCTGTGCCGTAC	4889
QY	4801	GCTATGACGGGTGACCGCGTATACAGGTTGCGCTGTGTGGAAGAAAGAACATTTGCCGTA	4860
Db	4890	GCAATGACAGGAGAAAGGATGCGCCGCTTAGTGTACACCAAGTTAAAGATGTGTGTT	4949
QY	4861	TGCTCATATTCCTGTTTCCGAAGTACAGATCACAGCGCTGTGCAAGCTACAGTGCAGC	4920
Db	4950	TGCTCATATTTTCCCTCTCCAAAATACCATGTAGATGGGGTGCAAGAGTTAAATGTGGAG	5009
QY	4921	AAACCAAGTCTGTTTTTACAGCGGTGTATCACCGGCTGTACACCCAGAGATTAAGCGGAA	4980
Db	5010	AAGGTTCTCTGTTTGACCCGACGCTTACTTCAAGTGTTAATCCGCGGAATATGCGCA	5069


```

RESULT 6
US-09-554-337-1
; Sequence 1, Application US/09554337
; Patent No. 6475780
; GENERAL INFORMATION:
; APPLICANT: Parrington, Mark
; APPLICANT: Li, Xiaomao
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: ALPHAVIRUS VECTORS FOR PARAMYXOVIRUS VACCINES
; FILE REFERENCE: 1038-1042 MIS
; CURRENT APPLICATION NUMBER: US/09/554,337
; CURRENT FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/065,791
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/CA98/01064
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 15538
; TYPE: DNA
; ORGANISM: respiratory syncytial virus
US-09-554-337-1

Query Match: 31.3%; Score 1564; DB 4; Length 15538;
Best Local Similarity 58.3%; Pred. Mismatches 2040; Indels 36; Gaps 8;
Matches 2004; Conservative 0;

QY 31 AGAATTCAAGTTAGACTTGAATGCTGACAGCCCGTATGTCAAGTGGTTACAGCGGAGCTTT 90
Db 2642 AAAGTGCATGTTGATATTGAGGCTGACAGCCCATTCATCAAGTCTTTGCAAGGCAATTT 2700
QY 91 CCACAATTGAGATGAAGCAAGGCGAGTCACTGCAATGACCATGCCAATGCCAGAGCG 150
Db 2702 CCGTGGTTCGAGGTGAGTGCATTCAGGTGCACACCAAAATGACCAATGCCAATGCCAGAGA 2760
QY 151 TTTTGGCATGTGGCAACAAAGCTCATTTGAGACGGAAGTGCACCGGAGCAAGTTATCTTG 210
Db 2762 TTTTGGCACTGGCGTCAACAAATGATGAGCAGGAGACTGCACAAAGACACACTCATCTTG 2820
QY 211 GACATTGGAAGTGGCCCGCTGCAGATGCATGCATTCACCGCTATCATTTGATCTGCG 270
Db 2822 GATATCGCATGTGGCTCTTCCAGAGAAATGATGTCTTAGCACAATACCACTGCGTATGC 2880
QY 271 CCTATGATTAAGCGCTGAAGACCCCGACAGACTCAACCGTATGACAAAGAAT-----T 324
Db 2882 CCTATGGCGAGCGAGAAAGACCCCGAAAGCGTGCATGACTACGAAAGAACTGGCAGCG 2940
QY 325 AAGAAAATGTCAATTACCGACAGAACTAGCTCTTAAAGCGGCGAGACCTGCTGAGTGC 384
Db 2942 GCCTTCGGGAAGGTGCTGGATAGAGATTCGACGAGAAATCAACCGACTGACACCGTGC 3000
QY 385 ATGTCAACACAGAGCGGAGAGCTCCATCTCTGTGTATGSCACAGAGCCGACCGTGTAG 444
Db 3002 ATGGCTTACGCAAGAGCTGTAATCTCTACCTTTTCTCTGCAATACAGAGTCAACGTTGCT 3060
QY 445 TACTTTGGAAGTGTGACAGTATACCAAGATGTGTACGAGTCCATGACACCGCATCAATC 504
Db 3062 ACGGAGCGCAAGGTGGCCGTATACAGAGACGTGTATGCTGTATCAATGACACCAATCGCTG 3120
QY 505 TACCAACCGGCGCTTAAAGAGTTTGAACAATTTACTGATATGCGTTTGAACAGACCCCT 564
Db 3122 TACCATTCAGGCGGATGAAAGGTGTCAAGACGCGGTATTTGGATTGGTTGACACCAACCCG 3180
QY 565 TTTATGTATCAAAAACATGCGAGGTTCCTACCTTACTTTCACACAGAACTGGGCTGTAGAG 624
Db 3182 TTTATGTTTGAACGCGCTAGACGCGCGGTATTCACACTTACGCAACACTGGGCGTACGAG 3240
QY 625 AGAGTATTGGAAGCACTTAACATTGGCTCGGTATCAAGATCTTCAGAGAGCAGGCTT 684
Db 3242 CAGGTGTACAGCGCCAGAGACATAGGATCTGTGTGAGCATCTCTTACATGAGGAGAACTC 3300
QY 685 GGAABAACCTCTCAATCTTTAGGAAGAAGGCTCAACCTACTAATATGATCATATTTCTCG 744

```

Db	3302	GGCAAACTGTCATTTCTCCGAGAGCAATTGAAACCTTCGCACACAGTATGTTCTCG	3361
Oy	745	GTGGGTTCAACATCTACACAGAAGATGATCACTGTTCACGTAGCTGCATCTTCCAAAC	804
Db	3362	GTAGGATCTCATTTGTACACTGAGAGCAAGAAAGCTACAGAGAGCTGGCACTTACCTCC	3422
Oy	805	GTGTTTCCACTTGAAGAAAGCTTAACTTCAACAGGTAGATGTGGACCATTTGTCAGCTGT	864
Db	3422	GTATTTCCACCTGAAAGGTAAACAAATCCCTTTTACCTGTAGTGTGCATATCCATGTCATGT	3481
Oy	865	GAAGGGTACGTCATCAAAAAGATACATCCAGCCAGACATATCCGTAAGCTTGAGAAC	924
Db	3482	GAAGGGTACGTAGTTAAGAAAATACATATGTGCCCCCGCTGTACGGTAAACGGTtagg	3541
Oy	925	TTGGCGTCCAAATGACATCCGAGGGGTTTCTTGAGTTTGCAAGTACACAGATACGCTGGC	984
Db	3542	TACGCCGTGACGTATACACGCGAGAGGATTCCTAATGTGCAGACACAGACACTGTCAA	3601
Oy	985	GGCGAGAGGGTTCTTTTGTGTGTGTATCGTATGTACCAAGCCACATTTGCGATCAGTG	1044
Db	3602	GGAAAAAGAGCTCATTCCTGTATGTACCTACCTCCCTCAACCATGTGTGATCAATG	3661
Oy	1045	ACAGGAGTTTCGGCACTGACGTTAGTGTGATACCGACAAAAACTATTTGGTGGGCT	1105
Db	3662	ACTGCGCATCTAGAGGACCGACGTCACACCGAGAGACGACGAAGTTGTATGTGGATYG	3721
Oy	1105	AACCAAAGGATTTGCTCATATGTATAGGACGCAAGAAATCTTAAACAAATCAGAACTAT	1164
Db	3722	AATCAGAGATATGTTGTACAGGAAGAACACAGGAAACCTTAAACGATBAAAGACTAT	3781
Oy	1165	CTATTTACAGTGTGCGCCAGGCGTTTTCAGAGTGGGCGCGTGAAACATCTGGCCGACTTG	1224
Db	3782	CTGCTTCGATTTGTGGCCGTGGCATTTTAGCAAGTGGGGAGAGAAATCAAGGACAGACTT	3841
Oy	1225	GACGACGAAGAAACTTAGGGGTGCGGAGGCGCACTTACTATATGGTCTGCTGGGCT	1284
Db	3842	GATATGAAAAACCTTGGGGTCCGAGAGGTCATCTTACTCTGCTGTGTGGGCA	3901
Oy	1285	TTCAAGACCCGAAATATCAATCCATCATCAAGAAGCGCTGTAGCAAAACAATTAAAGAA	1344
Db	3902	TTTTAAACGAGAAAGATGACACACATGTACAGAAACCAAGACCCAGACAAATAGTAGAG	3961
Oy	1345	GTACCTGCGCTTGTGACTCATTTGTGATTTCCAGCGCTTACCAAGCCAGGCTCGATATG	1404
Db	3962	GTGCTTCAAGGTTTAACTGGTGTGATCCGAGCGTATGTCTACAGGCGCTGGCAATC	4021
Oy	1405	GGCTTCCGCGGTAGGCTCAGCTGCTGCTTGAACCAACTGTCAAAACCGCACCGGCTATT	1464
Db	4022	CCAGTCAGATCACCATTTAAGATCTTTTGGCCAAAGAACCAAG---GAGATTTAATA	4078
Oy	1465	ACAAATGGCGATGTGAGACATCTGCGGGCTTAAAGACGAAGAGCTGAAGATGGGCTCA	1524
Db	4079	CTTGTTCTCGAGCGGTGTGTCAGCCAGGAGATGCTGAACAAGAGAAAGAGAGGTTGAG	4138
Oy	1525	GGCGAAGGATCAGAGAAGCCCTGCGCACCTTGTCTCTGTAATAAGAAAAAGAACGCTA	1584
Db	4139	GGCGAGGTGATBAGAAAGCTTACACACCTGCTCCCATCCGGCGCGCGAGAACGGGA	4198
Oy	1585	GAGCGAGAATGACCTCATTTATGCAAGAG-----GCAGAGCAGGTTAGCTGTGAGACA	1638
Db	4199	GTCTGTCAAGCTGACGTTGAAAGAACTAGAGATATCACGAGGTGACGAGGGGTGTGAAACA	4258
Oy	1639	CCAGAGAGACACATCAGAGGTGACAAAGTTACCCAGCGCAAGAGAAAGATTTGGCTTACGCT	1698
Db	4259	CTGTGACAGCGCTTGAAGTACACCGCACAGCGCAACGATCTACTAGGAAATTTACGTA	4318
Oy	1699	ATACTTTACCCACAGCGGATTTAGATATGTGAAAAATCGGCGTGTATCCACCAATTGGCG	1758
Db	4319	GTCTGTGTCCCGCAGACCGTGTCTAAGAGCTCCAAAGTTGGCCCCCGTGCACCTCTAGCA	4378
Oy	1759	GAACAGTACTGTATGACTCACAAAGTAGGCGCAGGAGATTCAAAGTGCAGCCATAC	1818

D 4379 GAGCAGGTGAAAAATATPAACATACGAGGAGCGCGGTTCACAGTGCAGCATAT 4438
Q 1819 CAGCGTAGGCTATTGTACGAGAGGAGCGCGGTTCCTGTTCAAGCTTCAGGCATATG 1878
D 4439 GACGGCAGGGCTCTACTACCAATGTGATCGGCATTCGGTCCCTGAGTTTCAAGCTTTG 4498
Q 1879 AGTGAAGCCGTACGATCGTTTTCACAGAGGAGGAGTTCTGTAACAGATACCTGACAC 1938
D 4499 AGCGAGAGCGGCATATGTGTACACGAAAGGAGTTCTGTCACAGGAAACTATACAT 4558
Q 1939 ATCCGAATCAACGAGAGAGCGCTTAACACTGACGAGAGTACTATTAAGACTGTAAAGCT 1998
D 4559 ATTCCTCTTCAAGCAGCGCTCGCTGACACCCAGCAGAGAACTACGAGAAAGTACAGACT 4618
Q 1999 CAGGACACAGCTGAGATAGTCTGTGATATGACGACGAGAAAGTGTAAAGCGAA 2058
D 4619 GAAGAATCTGACGCGAGTACGTGTTCGATGATTAATAAAATCTGTGTAAGAGAGAG 4678
Q 2059 GACGCAAGTCCCTTGTGCTTAACCGGTGATCTGTAGATCCACATTTTCAGAGTTTGG 2118
D 4679 GAACGTCGGGTTTGGTGTGGAGAGCTTAACCAACCCCGTTCCATGATTTGGCC 4738
Q 2119 TACGAGAGTCTAAGACACGACGACGACCTTCAAAAGTCCCAACCATGTGAGTCTAT 2178
D 4739 TACGAGAGGCTGAAGATCAGCGCTCGGACCATATTAAGACTACAGTAGTAGAGTCTT 4798
Q 2179 GAGATGCGAGGTGAGTAAATCTGAAATCATCAAAAGCGCTGTGATCAAAAGATCTG 2238
D 4799 GGGGTTCGGGATCAGCGAAGTCTGTATATTAAGACCTGTGACCAAAACAGATCTG 4858
Q 2239 GTTGTGATGCGAAGAAAGAACTGCGCAGAAATCATCAGAGATGTAAAGAGATGAGA 2298
D 4859 GTCAACAGCGGCAGAAAGAGAACTGCCAGGAATATGTAACGACGTGAACAGACACCGC 4918
Q 2299 CGATGTGATGTTGCTGCTAGAGCTGTGATTCAGTCTTCAATGCGGTTAAAGCACC 2358
D 4919 GGGAGGGGACAAAGTAGGAAACAGTGCATCTGCTAAACGGGTGTGCTGTGCC 4978
Q 2359 GTTAACCTCTGTACATTGATGAGGCAATTGCTGCACTGAGGAGCGCTGTGCACTG 2418
D 4979 GTGACATCTATATGTGAGCAGAGCTTTCCTGTGCAATCCGATCTCTGTGACCTTA 5038
Q 2419 ATTGCCATCTCAAACT---AAGAAAGTGTATTTGTCGGGGAGCCCAAACTATGCGGC 2475
D 5039 ATTCTCTGTTTAAACCTCGAGCAAGTGTGTTATCGGAGACCCCAACATATGCGGA 5098
Q 2476 TTCTTTAATGATGTGCTTGAAGTACATTTTAACTGACATATGACATGAGTGTAC 2535
D 5099 TTCTTCAATATGATGACGCTTAAGGTAACTTCAACCAACATCTGCACTGAAGTATGT 5158
Q 2536 CATAAAGCATCTCTAGAGAGTGCACACAGACTGTAAACGCGCATGCTCCACGCTTC 2595
D 5159 CATAAAGTATATCAAGACGTTGACGCTGCACTGACGCGCATGCTTCTACGTTGAC 5218
Q 2596 TACGACAGCGAATGAAGACGTTAAACCATGTGCTGATTAATCATCATGATATACCA 2655
D 5219 TACGAGAGCAAGATGCGCAGACCAACCGTGCAAAACCATATATATGACACCA 5278
Q 2656 GGAACACAAAGCCGACAAAGATGATGTATTTCACTGTTTCAGAGATGGGTAAA 2715
D 5279 GGAAGACCAAGCCCAAGCAGAGACATGCTGTATTAATCATCTTCGAGGCTGGCAAG 5338
Q 2716 CAGCTACGATTTGATCAAAATCAGAAATCATGATGCGGCTGCACTGCCAGGACTT 2775
D 5339 CAGCTGAGTTGACTACCTGTGACAGAACTGACGAGACGACATCTCAGGCGCTC 5398
Q 2776 ACGCGAAAGCGTTTATGCTGTGAGTAAAGTCAACGAGAAATCATCTACTCGAG 2835
D 5399 ACCCGCAAGAGGATATACGCGTAAAGGACGAAAGTGAATGAATTCCTTGTATGCCCT 5458
Q 2836 ACTTCTGACAGTGAACGTTTACTTACACGACAGAAAAACGATTTCTGAGAGAG 2895
D 5459 GCGTCGAGCAGTGAATGTACTGTGACGCGCACTGAGATAGGCTGTGTGAAAAAG 5518

Q 2896 CTAGCTGATCCCTGATTAAGACACTTACAGCTAAATATATCCGGGATTTCAACGGCT 2955
D 5519 CTGGCCGGCATCTCCCTGATTTAAGTCTTATCAAACTTCCACAGGTAATTTACGCC 5578
Q 2956 TCATTTGACGACTGCGACGCGCAACAGACCCATTAATGACGCGTTCTTGATMAAGCG 3015
D 5579 ACATTTGAAGATATGCAAGAAACAGCAAAATATTAAGAGTATTAAGAGACGGCT 5638
Q 3016 CAGACAGTATGTTTCCAGAAATAGTAAAGTCTGTGGCGGAAGCTTTAGAACCA 3075
D 5639 GCGCTGTGACGCGTTCCAGAACCAAGCAAGCTGTGTGGCGGAAGACCTGTGCTCT 5698
Q 3076 GTCTTGCACAGCGCAATTTGTGCTGACAGACAGCAGTGGGAGAC---GTTCACCCA 3132
D 5699 GTCTGACACTGCGCAATTCAGATTGACAGCAGAGAGTGTGACACCATATATACAGCA 5758
Q 3133 TTCAAGCATGACAGAGGTACTGACCTGAATGCACTGAACCTTTTGTGACACGAGTTC 3192
D 5759 TTTAAGAGAGACAGAGCTTACTCTCCAGTGTGGCTTGAATGAATTTGACCAACAGTAC 5818
Q 3193 TTTGAGATGACCTGACAGTGGTTATTTCCGCTCTTACCGTGCATCTTACTTACAGG 3252
D 5819 TATGAGTTGACCTGACAGTGGCTGTGTTTGTGCCCCGAAGTGTCTGTATTTACAG 5878
Q 3253 GATCAGCACTGGGATTACTGCGCAGGAGAAACATGTATGGCTTAATAGAGAGTACGA 3312
D 5879 AACCAACACTGGGATTAACAGACCTGTGTGAAGATGTATGATTCATATGCCGCAACGCT 5938
Q 3313 AAGAGTTGTACCGCGATATCCGTGCATCAAAAGCGGTGACACAGCAGGAGTACT 3372
D 5939 GCCAGGCTGGAAGCTAGACATACCTCTTGAAGGGGAGTGTGCTACGCGGCAACAGCA 5998
Q 3373 GATTTAAGATATATACCATCAAGAGTACTCTTCCAACTTAATGTGTTCCTTAAT 3432
D 5999 GTTATCCAGAAAGAAAAATCAACGCTTGTGTGTGTGACAAATGTATTTCTATCAAC 6058
Q 3433 CGCGGTTGCCCCACTGTTGATGTTTGAACCAAAAGAGAGGTAACAATGATCACAGC 3492
D 6059 CGCAGGCTCGCAGCGCTGCTGTGTGATGACAAAGACGTTAAAGGCAATGAGTTGAG 6118
Q 3493 GGATTCCTATTAAGATGAAGGCAATCTGTGTGTGATTCGGCATCTTATCAGATT 3552
D 6119 TGCTGTGATTAATGAAGAGGTAACAGTCTCTGTGTGATGATCAACCTGGCT 6178
Q 3553 CCAGGAGAAAGTATGATGCTCAATGGCTCAATGGCCCTAATC-----CATCAG 3603
D 6179 TTGCTTGACGCGAGGTCATCTGTGTGTCAACCGTGAATGACAGGCGCGGATAGTGC 6238
Q 3604 TGTATCTGATTTGGGAATACCTAGACATGTCCGTTAAATATGACATATCTTTGCAAT 3653
D 6239 TACGACTAAGTTTAGACATGCGCGCTGACGCGCGAGTTGCACTGTGCTTTGTGAC 6298
Q 3664 GTTAGAACCCCGTACAGAAACATCACTACCAAGTGTGAGATCAACGCTATCACAC 3723
D 6299 ATTCAACGGAATTCAAATTCACCACTACAGCAGTGTGTGACACAGCCATATACCTG 6358
Q 3724 AGCATGTCAAGTGTAAAGCTGTGTCAACCTGAACCTGCGGGAACATGTGTGCTATA 3783
D 6359 CAGATGTCTTGGGAGATGTGCTACGACTGTAAACCCGCGGCACT---TATAGAGA 6415
Q 3784 GGGTATGGCTTGTGTATTCGCGCAACCGAGAAATTCATCACTGCGGTGTGACGCTCAAT 3843
D 6416 GCTTAGGATATGCGCATTAATATGAGGAAGCGTTGTTCTCTTAAAGAGAAATTC 6475
Q 3844 AGGTTTACCGGTGTCTGTACGCTTAAAGAACCTGCGGAATACTGAGGTTCTTCTG 3903
D 6476 TCGTCTCAAGAGTGTGCGCGCATTTGTGTCAACGAAATACAGAAAGTGTCTTCTG 6535
Q 3904 TTCTTGGCAGAGCAACGCGCAACACACATGACAGAGACAGACTCGGTGTAGTCTT 3953
D 6536 TTCTCCAACTTTTGAACGAGAAAGAACCTCTTACGTTACACGAGATGAATACCAAGCTG 6595

3964 GACAACTATTCAGAGGTCAACCAAGTACGAGGAGGAGGAGTCCACCGTACAGAGT 4023
Db AGTCCGATATGCGCGAGAAACCAATGACACGCGCGGATGTCACCACTTACAGAGT 6655
4024 ATCAGAGGATATAGCAAGAGCGGTGACCAAGTATGCTTAATGCTGCTAATAGCAA 4083
Db AAGAGAGAGATATAGCCACGTCACAGAGCGGCTGTGTTAGCGAGTAAACCGCT 6715
4084 GGTCAACAGAGTCCGAGTGTGCGGTGCACTGTACCCGAAATGCGCGCTCTTTGAT 4143
Db GGAATGTAGAGGAGATGCGGTATGAGGCGCGGTGCGGAAATGCGCGCTTTAG 6775
4144 AGACAGCAATATGCTGTGGAGCGGTAGACTTTGTAACAGAAACCG--CTCATATA 4200
Db GGAGCAGCAACACAGTGGGCAATTAACATGATGTGGGCTGTAACCCGCTCATC 6835
4201 CATGTATAGAGCCCAATTTTCTAAGATGCGGAAACCGGAGGAGGAGCTTAAGTCCGA 4260
Db CACGCTGAGCGCTTAATTTCTCTGCAAGTGAAGCGGAGGAGGAGCGCAATTTGCC 6895
4261 GCTGCTCATGAGATAGCGTCCATGCTCAACGCTGAGCGGATTAACAAAATATCAGTA 4320
Db GCTGCTACCGGAGAGTGGCGCCCAAGTAAACAGCTGTCACTGACGAGCTAGCCATC 6955
4321 CCGTACTGTCAACCGGATCTATTCTGTGGCAAGATGAGTATGATCATTTGAT 4380
Db CCGCTGTGTCCAGAGAGTGTTCAGCGCGGAGAGATAGGCTGAGCAATCCCTCAAC 7015
4381 CACCTGTCTATGCTTTGAGCATAGAGTCCGATGTCACTATATTTGCTTGATAA 4440
Db CATCTATTCACAGCAATGAGCGCAAGCGGTGAGTCACTATCTACGAGAGACAA 7075
4441 CAATGGAGACGAGATATCGAGCCATTCACCCCAAGAAAGCGTAATTCGAT 4500
Db AGTTGGAGAGAAATCCAGAAAGCCATTCAGAGAGCGGCTGTGATGCTCAAT 7135
4501 GATGCAAGCAAGTATGATGATGCTTGTGAGGTCACCCAAACAGCTTTTGGCAGGC 4560
Db GATGACGTGAGACTACCAAGACTTGTGTAGAGTGCACCGGACAGCGCTGTGGGT 7195
4561 AGACCAAGTACTCCGTCAATGAGGCAAGTTGTATTCATACCTGGAAGTACAGATTC 4620
Db CGTAAAGGCTACAGTACCACTGACGCGTGTGCTGTACTTTGAAAGTACGAAATTC 7255
4621 CATCAACCGCCAGAGCATTTGCCGAAATTCATGCAATGTGGCCCAACAAATCTGAGGT 4680
Db AACCGAGCTGCTATTTGATGTCAGAGATCTGACGTTGTGGCCAGACTGCAAGAGGA 7315
4681 AATGAGCAATTTGCTGTATCATCTGTGGGAGAGTATGTCAGATCCGCTCCAAATGC 4740
Db AACGAAAGATATGCTATAGCGGCTGTGGCGAAACATGGAACATGATCCAAATGT 7375
4741 CCAATGAGAGTCAAGAGCGCTGTCTCACTCAACACTTCATGCTGCTGTATTAAC 4800
Db CCGGGAACGATTCGATTCATCAACACTCCCAAGACAGTGCCTGTGCGCGTAC 7435
4801 GCTATGAGGCTGAGCGGATATCAAGTGTGCGCTGTGGAAGAAAGCAAGTGTCCCTTA 4860
Db GCAATGAGAGCAAGAGATGCGCGCTGTAGTACACCAAGTTAAAGCATGTGTGT 7495
4861 TGCTCATCATTCCTGTCGCGAAGTACAGATCAAGCGGTGCAAGATCAAGTACAGCAGC 4920
Db TGCTCATCTTTTCCCTCCGAAATACCATGTAGATGGGTGCAAGATGAAGTGTCCAG 7555
4921 AAACCAAGTCTGTTTCAAGCGCTGTACCAACGCGTGTACACCCAGAGAAAGTACGCGAA 4980
Db AAGTTCTCTGTTGACCGCGAGTACTTCAAGTGTAGTGTCCGCGAAGATATCCGCA 7615

RESULT 7
US-09-554-337-4
; Sequence 4, Application US/09554337
; Patent No. 6475780

GENERAL INFORMATION:
; APPLICANT: Paringdon, Mark
; APPLICANT: Li, Xiaomao
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: ALPHAVIRUS VECTORS FOR PARAMYXOVIRUS VACCINES
; FILE REFERENCE: 1038-1042 MIS
; CURRENT APPLICATION NUMBER: US/09/554.337
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/065,791
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/CA98/01064
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 8100
; TYPE: DNA
; ORGANISM: respiratory syncytial virus
US-09-554-337-4
Query Match 29.5%; Score 1473.8; DB 4; Length 8100;
Beet Local Similarity 57.9%; Pred. No. 0; Mismatches 1982; Indels 36; Gaps 8;
Matches 2779; Conservative 0;
214 ATTGAAAGTGGCGCCGTCAGACATGACATTCATCCATCCCTATCATTTATCTGCCCT 273
Db 1 ATCGCAGATGGCGCTTCACAGAGATGATGTCTACGCAAAATACCATCTGCGTATGCCCT 60
214 ATGATAGCGCTGAAGACCCGAGACAGACTACACGATACACGATATGCAAGAAAGCTTAAGAAAA-- 331
Db 61 ATGCGCAGCGGAGAAAGCCCGAAAGGCTCATATGCTACGAAAGAACTGCGACGCGCC 120
332 ----GTGACATTACCGACAAAGACATGACCTCTAAGCGGAGAGCTCTGGAAGTCAAT 387
Db 121 TCCGGAGAGTGTCTGATAGAGATGCGAGAAATACCGAGCTTCAAGCGCTCAG 180
388 TCAACACAGAGAGAGAGACTTCATCTGTATATGACAGAGAGCCGAGTGTAGTAC 447
Db 181 GCTACCGCAGAGCGGTGATCTCTATCTTTGCTGCTGATACAGAGTACCGTGTGTACG 240
448 TTGGAAGTGTACAGATATACCAAGATGTGTACGAGTCCATGACCGACATCAATCTAC 507
Db 241 GCAGCGAAGTGGCGGTATACAGAGCGTATGTCTGTATGCTATGACCAACATCCGCTGAC 300
508 CACCAAGCGGCTTAAGAGGTTAGACAAATTTAGTGAAGCTTTGACAGACCCCTTT 567
Db 301 CATAGCGGATGAAGGTTGAGAAAGCGGCTATTTGATGATGATTTGAGACCAACCCGTTT 360
568 ATGTACAAATCATGAGCGAGTTCCTTCACTTCAACAGAGAGTGGCTGACAGAGA 627
Db 361 ATGTTTACGCGCTAGACAGCGCGCTATCAACTTACGCAACAACTGGGCGAGAGCAG 420
628 GTATTTGAAGACGTAACTTGGCTGTGTAATCTAGATCTTCAAGAGAGAGCGTTTGA 687
Db 421 GTGTACAGGCGCAGAGACATATGACTGTGTCAGATCTTGAATGAGGAAAGTACCGG 480
688 AAATCTCAATCTTGAAGAAAGAGGCTCAACCTACTAATTAAGATATATTTCTCGGT 747
Db 481 AAATGCTCATCTTCCGGAAGCAATTTGAACCTTGCGACACAGATATGTTCTCGGTA 540
748 GGTTCACAACTCTACAGAGATAGATCACTGTTAGTAGTGGCATCTTCCAAAGCG 807
Db 541 GGAATCTACATTTACCTGAGAGCAAGAGCTTACTGAGAGCTGCACTTACCTCCGTA 600
808 TTCACTTGAAGAAAGTCTAATCTCAACAGTATGATGTGAGACCATTTGACGTGAA 867
Db 601 TTCACTTGAAGAAAGTCAATCTTACTGTATGATGTGATGATGATGATGATGATGAT 660
868 GGTATCTCATCAAAAGATATAGATCAAGCCAGAGCTATACGTTAAAGTTGAGAACTTG 927
Db 661 GGTATCTGATTAAGAAATCATATGTGCTCCGCGCTGTAGCGTAAACGATGAGGTAC 720
928 GGTTCACAAATGATCGGAGGTTTCTTGTGTTGCAAAAGTACAGATACCGTCCGCGC 987

Db 721 GCCGTGACGTATCAACGCGAGAGGATTCCTAGTGTGCAAGACCAACAGACATCTGTCAAAAGGA 780
Qy 988 GAGAGGCTTTCTTTTGTGTGTGTATGATATACAGACCACTTTGGATCAATATGACA 1047
Db 781 GAAAGAGTCTATTCCTCTGTATGCACTTACCTCCCTCAACCATCTGTGATCAATATGACT 840
Qy 1048 GAGATTCGTGCACATGACGTAGTGTGATGACGACCAAAATAATTGGTGGGCTCAAC 1107
Db 841 GGCACTACTAGGACCGACGTATCAACCGAGAGACGACAGAAATTGTATGGGATTTGAAT 900
Qy 1108 CAAGAGATTGTCTCAATGTATGAGACGCAAAAGAAATATCAACATGACAGAACTATCTA 1167
Db 901 CAGAGATATGTTGTGAACGGAAGAACACAGGAAACATTAACGATGAAGAACTATCTG 960
Qy 1168 TTACCACTAGTGTCCCGACGCGCTTTCCAGGTGGGCGGTGAACATCTGTCCGACTTGGAC 1227
Db 961 CTTCGATTTGTGGCGGTTCGCACTTTAGCAAGTGGGCGAGGGAATCAAGGCGACACTTGTAT 1020
Qy 1228 GACGAGAAAGAACTAGGGGTGCGGGAAGGCACTCTTACTATAGGGGTGCTGGGCTTTC 1287
Db 1021 GATAAAAAAGCTCTGGGTGTCCGAGAGAGTCACTTACTTGTCTGTCTGTGTGGCATTTT 1080
Qy 1288 AAGACCCAGAAATCAATCCATCTCAAGAGAGCTGTGTAAGCAACATTTAGAAAGTA 1347
Db 1081 AAAACGAGAGATGCAACACCATGTACAAAGAACCAAGACCAACGAAATGTGAAGTG 1140
Qy 1348 CTTGCCCTTTGATCTATTTGTATTCACGCTTACAGCCACAGGCTGATATGGGC 1407
Db 1141 CTTTCAAGATTAACTGTTGTGTATCCGAGCCTATGTGTCTAAGGCTGTGCAATCCCA 1200
Qy 1408 TTCCGCGTAGGCTCAAGCTCTGCTTGAACCAACTGTCAACCCGACCGGCTATTACA 1467
Db 1201 GTCAAGATCAAGCATTTAGATCTTTTGGCCAAAGAACCAAGCCAGA---GTAAATACCT 1257
Qy 1468 ATGGCCGATGTGAGCATCTGCGTGTCTTACAGCAAGAGCTGAAGAGTGGCTGCAGCG 1527
Db 1258 GTTCTCGACGCGGTGTCAGCCAGAGGATGTCTGAACAAAGAGAGAGAGGTTGAGGCC 1317
Qy 1528 GAAAGATCAGAGAAAGCCCTGACCCCTTGTCTCTGTAATATGAAAAAGAACCTGTAGAG 1587
Db 1318 GAGTGAATAGAGAGCTTTCACACCCCTCTCTCCCATGCGCGCGGAGAGAGGAGTGC 1377
Qy 1588 GCAGAGTAGACCTCATTATGCAAGAG-----GAGAGAGAGGATGTGAGACACCA 1641
Db 1378 GTCAAGTCAAGCTTGAAGAACTAGATATCAACGAGGTGCAAGGCTGTGAAACACCT 1437
Qy 1642 CGAGAGACATCAGGCTGACAAAGTTACCCAGGCGAAGAGAAAGATTGGTCTTACGCTATA 1701
Db 1438 CGCAGCGGTTGAAGTCAACCGCACAGCGCAACGACTACTAGGAAATTTACGTAGTT 1497
Qy 1702 CTTTCAACCCCGAGGGGATTTGAATATGTAATAAACTGGCGTATCAACCTTGGCGGA 1761
Db 1498 CTGTCCCGAGACCGTGTCTCAAGAGCTCAAGATTGGCCCCCGGTGACCCCTCTAGCAGAG 1557
Qy 1762 CAATACTGTATGATGATCACAAAAGGTAGGGCAGGAGATCAAAAGTGAGCCATPACAC 1821
Db 1558 CAGGTAAATATATACATACATACGAGAGGAGCGGCGGTATCCAGGTGACGATATGAC 1617
Qy 1822 GGTAAAGTATTTATCAAGAAAGGACGCGGCTCCCTGTTCAGACTTCCAGGCAATTAGT 1881
Db 1618 GCGAGGGGTCTACTACATGATGATTCGGCATTCGGGCTCCCTGAGTTTCAAGCTTTGAGC 1677
Qy 1882 GAGAGCGCTAGATCGTTTCAAGAGAGGAGATTCTGAACAGATACCTGACCAATC 1941
Db 1678 GAGAGCGCACTATGATGTATCAAGAAAGGAGATTCTGCAACAGAAACTATATACATAT 1737
Qy 1942 GCATATCAAGAGAGCGCTAAACATGACGAGAGATCTATAAGACTATAAGACTAG 2001
Db 1738 GCGCTTACAGGACCGTGTGCTAACAACGACGAGAGAACTACGAGAAAGTACAGACTGAA 1797
Qy 2002 GACACAGACTCAGATATGCTTGTGATATTTGACCAAGAAAGTGTGTTAAGCGAGAAC 2061

Db 1798 AGAAGTACCGCCGAGTACGTGTTCGACGTATATAAAAATGCTGCTCAAGAGAGGAA 1857
Qy 2062 GCAGGTCCCTTGTGCCCTAACCGGATCTGTATGATCCACCATTTACGAGTTTGGCTAC 2121
Db 1858 GCGTGGGTTTGGTGTGTGGAGACTTAACCAACCCCGTTCCATGATATTCGCTTAC 1917
Qy 2122 GAGAGTCTCAAGACACGACGACGACGACCTACAAAGTCCCAACCATCGAGTCTATGGA 2181
Db 1918 GAAGGGCTGAAGATCAGCGCGTGGGACCATATTAAGCTACAGATAGAGAGTCTTTGGG 1977
Qy 2182 GTGCCAGTTTCAAGTAAATCTGGAATCATAAAAGCGCTGTGACTAAGAAAGTCTGTT 2241
Db 1978 GTTCCGGGATCAGGCAAGTCTGTATATTAAGGCTCTGTGACCAACAGATCTGTC 2037
Qy 2242 GTGAGTCCGAAGAAAGAACTGGCGGAATATCATCGAGGATGATAGAGAGATGAGAGCT 2301
Db 2038 ACCAGCGGCAAGAGAGAACTGCGAGAAATGTTAAACGATGAAAGAGACCGGCGG 2097
Qy 2302 ATGATGTTCTGCTAGACCTGTGATTCAGTGTCTTAAATGGGTTAAGCACCCCGTT 2361
Db 2098 AAGGGAACAGTAGGAGAAACAGTGAATCATCTGCTTAAACGGGTGTGCTGCGCGTG 2157
Qy 2362 AACCTGTATCATTTGATGAGGCAATTTGCTGCCATGACGAGGAGCTGTGCACTGAT 2421
Db 2158 GACATCTTATATGTGACGAGGCTTTGCTGCAATCCGTAATCTGTGCTGGCCCTAAT 2217
Qy 2422 GCCATGTCAAACCT---AAGAAAGTGTATGTGTGGGGGACCCAAACATGGCGCTC 2478
Db 2218 GCTCTGTAAACCTCGAGCAAGTGTGTATGCGAGAGCCCAAGCAATGCGAATTC 2277
Qy 2479 TTTAACTGATGTGCTGAAAGTACATTTTAACCATGATATGACATGAGTATACAT 2538
Db 2278 TTCAATATGATGACGCTTAAAGTGAATCTTACACCAACATCTGCACTGAAGTATGTCAT 2337
Qy 2539 AAAAGCATCTTATAGAGGTGACACAGACTGTAAACCGCATCTGCTCAGCGCTTTCTAC 2598
Db 2338 AAAAGTATATCCAGACCTTGCAGCGCTTCAGTCAACGCGCATCTGTGTACGTTGCACTAC 2397
Qy 2599 GACAAAGGATGAAAGAGGTTAACCCCATGTGCTGATTAATTCATATAGATACAGAGG 2658
Db 2398 GAGAGCAAGATGCGGACGACCAACCGGTGCAACAAACCATTAATCATATAGACACCAAGGA 2457
Qy 2659 ACCCAAGAGCGGACAAAGATGATCTGATTTCTAACCCTGTTTACAGAGAGGGTGAAAAG 2718
Db 2458 CAGACCAAGCCCAAGACGAGAGATCTGTATTAACATGCTTCCGAGGCTGGGCAAGACAG 2517
Qy 2719 CTACAGATTGATCAAAATTAACGAAATCATGATCTGCGCTGTATGCAAGGACTTACG 2778
Db 2518 CTGCAATTGATCACTGCTGACACGAAATCATGACACAGCATCTCAGGCGCTCACCC 2577
Qy 2779 CGGAAAGCGTTTATGTGTCAAGGTACAAAGTCAAGAGAAATCACTCTACTGCGACACT 2838
Db 2578 CGCAAGGGGTATACGCTGTAAGGCAAGAGTGAATATAATCTTGTATGCCCCCTGCG 2637
Qy 2839 TCTGAGCAGTGAACGTGTACTTACACGACAGAAAAAGCAATGCTGTGGAAGACGCTA 2898
Db 2638 TCGAGACAGTGAATGATGACTGTGACCGGCATGAGGATAGGCTGTGTGGAAGAACGCTG 2697
Qy 2899 GCTGTGATCTCTGATTAAGACACTTACAGCTAATATATCCCGGGGATTTTCAAGGCTTCA 2958
Db 2698 GCGCGCATCTCTGATTAAGGCTCTATCAAACTTCCACAGGGGTAACTTTACCGGCCACA 2757
Qy 2959 TTGAGACATGAGCAGCGCGGAACAGAGGCAATATGACAGCGGCTCTTGAATAGCGCGAG 3018
Db 2758 TTGGAAGAAATGCGCAAGAAACAGCAAAATATATGAAGTATTAAGAACCGGCTGCG 2817
Qy 3019 ACAAGTATGTGTTCCAGAAATTAAGGTGACGTCTGTGCGGAGAGGCTTTAGAGCCAGTC 3078
Db 2818 CCGTGTGACCGGTTCCAGAAACAAAGCAAGTGTGTGGGGAAGAACCTGTGTGCTGTC 2877
Qy 3079 TTGGCCACGCGCAACATTTGTGCTGACGAGACAGCATGTGGAGAC---GTTGACCCCATTC 3135
Db 2878 CTGGAACACTGCGGAATCAAGATTGACAGAGAGATGGAGACCATATATTAACAGCATTT 2937

3136 AAGCATGACAGACCGTACTACCTGAAATGSCACTGAACCTCTTTGACAGGTTCTTT 3195
2938 AAGGAGACAGAGCTTACTCTCCAGTGTGCTTGAATGAATTTGACCAAGTACTAT 2997
3136 GAGATGACCTGACAGTGGGTTATTTCCCTCTCAACCGTCACTTAATTACAGGAT 3255
2998 GAGATGACCTGACAGTGGGTTATTTCCCTCTCAACCGTCACTTAATTACAGGAT 3057
3256 CAGCATGGGATTAATCCGACAGGAAAGATGTAATGGCTTAATGAAAGTGAAG 3315
3058 AACCATGGGATTAACAGACTGATGAGAGATGTAATGATTCACCAACAGCTGCC 3117
3316 GAGTGTCAAGCGCATATCCGTGATCAAAAAGCGTTGACACAGAGGATGAT 3375
3118 AGGCTGGAAGTACATCTCTTCCGAAAGGAGCTGATACGAGCAAGGAGAT 3177
3376 ATAGGAATTAATCACTCAAGGATCTCTCAACATTAATGTGTTCCATTAATCCG 3435
3178 ATGCAAGAAAGAAATCCACCGCTTCTGTGCTGACATGTAATCTCAACCGC 3237
3436 CCGTTGCCCACTCTGTTGATCTGTGACCAAGAGAGGATCAACTATCAAGCGA 3495
3238 AGGCTGCGCAGCGCTGTGCTGATGACAGACGTTAAAGGAGTGGGTTAGTGG 3297
3496 TTCTATCTAAGATGAAGGCAAAATCTGTGTGATGCGCGATCTATCAGATCCA 3555
3298 CTGTGATTAAGTAAAGGAGGATACACGCTCTGTGTGATGATGATCACTGGCTTG 3357
3556 GGGAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3606
3358 CTGCAAGCAGGATCACTGTGTTGTCACCGCTGAATGTCAAGGCGCGATAGTGTAC 3417
3607 GATCTGATTTGGGAATTAAGTACCTGATGCTGCTGATTAATGATTAATCTTTGCTCAATGTT 3666
3418 GACCTGATTTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3477
3657 AGGACCCGCTACAGGAACCATCACTACCAAGTGTGAGATGATGATGATGATGATGAT 3726
3478 CACACGGAATTAAGATTCACACATCAACAGCTGTGTGATGATGATGATGATGATGAT 3537
3727 ATGCTACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3786
3538 ATGCTTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3594
3787 TATGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3846
3595 TACGATACGCTGATTAATTAATCAAGCAGCGCTGTTCTCTTAAGCAAGATGTTCTG 3654
3847 TTACCCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3906
3655 TCTGCAAGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3714
3907 TTGCGCAAGACAGCAAGCAACCAACATCAAGTACAGAGAGATGATGATGATGATGATGAT 3966
3715 TCCAACTTTGACAGGAAAGAGACCTCTAGCTACACATGATTAATCAAGCTGATGAT 3774
3967 AACATCTATCAAGGCTCAACAGGATCAAGGAGGAGGATGATGATGATGATGATGATGAT 4026
3775 GCGGTGTGCTGCGAGAAAGCAGTACAGCGCGGTGTGATGATGATGATGATGATGATGAT 3834
4027 AGAGTGAATTAAGCAAGGCTGATCAAGCTATGCTTAATGCTTAATGCAAGATGATGAT 4086
3835 AGAGCAAGATTAAGCAAGGCTGATCAAGGCTGATGCTTAATGCTTAATGCAAGATGATGAT 3894
4087 CAACAGGTTCCGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 4146
3895 ACTGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3954
4147 CAGCAATATGCTGTGAGGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 4203
3955 GCAGCAACCAAGTGGGCAATTAATCAAGTATGATGATGATGATGATGATGATGATGATGAT 4014

4204 GCTGATGAGACCAATTTTCTTAAGATGCGGAAACCGAGGCGACCTTAAGCTCCAGCT 4263
4015 GCTGATGAGACCAATTTTCTTCCACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4074
4264 GCTTACATGAGCATAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4323
4075 GCTTACCGGAGCAGTGGCGCGCGAGATGAACAGACTGCTACATGAGAGGCTGATGATGATGAT 4134
4324 CTACTGTCACCGGATCTATTTCTGTTGCGAAAGATGATGATGATGATGATGATGATGATGAT 4383
4135 CTGCTGTCACAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4194
4384 CTGCTGTCACAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4443
4195 CTATTCACAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4254
4504 GACCAAGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4563
4315 GACGTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4374
4564 CAGGTTACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4623
4375 AAGGCTTACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4434
4624 CAGACCGCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4683
4435 CAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4494
4684 GAGCAGATTTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4743
4495 GAAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4554
4744 GTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4803
4555 GTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4614
4804 ATGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4863
4615 ATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4674
4864 TCATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4923
4675 TCATCTTTTCCCTCCCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4734
4924 CAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4980
4735 GTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4791

RESULT 8
US-08-446-932-1
Sequence 1, Application US/08446932
Patent No. 5639650
GENERAL INFORMATION:
APPLICANT: Johnston, Robert E.
APPLICANT: Simpson, Dennis
APPLICANT: Davis, Nancy L.
TITLE OF INVENTION: cDNA Clone for South African
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenneth D. Sibley
STREET: Post Office Drawer 34009
CITY: Charlotte
STATE: NC
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,932
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-118
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 881-3140
TELEFAX: (919) 881-3175
TELEX: 575102
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11663 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEICAL: NO
US-08-446-932-1

```

Query Match 26.7%; Score 1333.2; DB 1; Length 11663;

Best Local Similarity 56.1%; Pred. No. 0; Mismatches 2128; Indels 80; Gaps 13;

Matches 2820; Conservative 0;

```

QY 23 ATATGAAAGAAATTCAGCTTACCTGATGCTGACAGCCCTGATGCTCAAGTCTTACAGC 82
DB 64 AGAAGCAGATGTTACGTAGACGTAGACCTCAGAGTCCCTTGTGTGCTGCACTGCAAA 123
QY 83 GGACGTTTCCAAATTTGAGATGCAAGCAGGCAAGTCACTGCAATGACCTCCAAATG 142
DB 124 AGACCTTCCCGCAATTTGAGATGAGTACAGCAGAGTCACTCCAAATGACCATCTAATG 183
QY 143 CCAGAGGTTTGGCAGTGGCAAAAGCTTATGAGAGGAGTGCAGCCGGAGCCAG 202
DB 184 CCAGAGGTTTGGCAGTGGCAAAAGCTTATGAGAGGAGTGCAGCCGGAGCCAG 243
QY 203 TTATCTTGAACATTTGAGAGTGGCCCGCTCAGACATGACATTTCCAAATCAGCTATCAT 262
DB 244 CGATTTTGGACATGAGGAGCGGCAAGGCTGTAGATGTTTCCGAGAGCAGATACAT 303
QY 263 GTATCTGCTTATATAGCCCTGAAGACCCGACAGACTACAAAGTATGCAAGAAAG 322
DB 304 GCGTTTCCCGCATGATGTCAGAAAGCCGGAACCGCATGATGAAATATGCCAGCAAC 363
QY 323 TTAAAGAAA-----GTGACATTAACCGAAGAAATGAGCTTAAGGCGGAGAGCTGC 376
DB 364 TGGCGGAAAAGCATGTAGATTAACAAAGAACTTGATAGAAAGTCAAGAGCTCC 423
QY 377 TGAAGTATGTCAACACAGAGCCAGAGCTCCATCTGTGTATGACACAGAGCCCA 436
DB 424 GGACCGTACTTGAATACACCGGATGCTGAAGCCATCATCTGCTTCCAGACAGATGTA 483
QY 437 CGTGTAGTACTTTGAAAGTGTAGACATATACAAAGTGTGACGAGTCCATCAGCA 496
DB 484 CTTGCAACACGCGTGGCAGATGCTCCGTCATGACAGAGCTGTACA---TCAACCTCCG 540
QY 497 CATTAATCTACACAGGCGCTTAAAGAGTTAGAGCAATTTATGATAGCTTTTACA 556
DB 541 GAATCTATTACACAGGCTTAAAGAGCGTGGACCTGTGATGATGAGTTGGCTTCA 600
QY 557 CGACCCCTTTTATGACAAAACATGCGAGGTTCTTACCTTATACACAGAACTGGG 616
DB 601 CCACCCAGTCTATTTCTCGGCTATGCGAGTGTGTAACCTGACATACACCAACTGGG 660
QY 617 CTGACAGAGATTTGAGAGCAGTAAATTTGCTGCTGTAATCTCAGATTTTCAAGAGA 676
DB 661 CCGACGAAAAGTCTCTGAGGCGGCTAAATGAGCTGACGACAAAGCTGAGTGAAG 720

```

```

QY 677 GCAGGCTTGGAAAACTCTCAATCTTATGAGAAAGAGGCTCCAACTTAAATAGATCA 736
DB 721 GCAGGACAGAAAGTCTTCAATATGAGAAAGAGTGAAGCCCGGCTACGGGTTT 780
QY 737 TATTTGGTTGGTTTCAACAATCTACAGAAAGATGATCACTGTTACGTAGCTGGCATC 796
DB 781 ATTCTCGTTGATTCACACTTTTACCAAGAACAGAGCCAGCTTGCAGAGCTGCATC 840
QY 797 TTCCAAAGCTTCCATTTAAAGAAAGTCACTTACAGAGTATGAGGAGCCATTG 856
DB 841 TTCCATCGTGTTCATTTAAAGAAAGAGTGTACCTTGCCTGTGATCACTGG 900
QY 857 TCAGCTGTGAAGGATGATCATCAAAAGATTAAGATCAGCCAGACTATACGTTAAG 916
DB 901 TGAGCTGCAAGGCTAGTATGTAAGAAATCAGATCACTCCCGGATCACGGAGAAA 960
QY 917 TTGAGAACTTGGCGTCCAAATGATGCGAGGCTTTCTTGAATTGCAAAAGTCA 976
DB 961 CCGTGGATACGCGGTAAACAAATGAGGAGGCTTCTGTATGCAAAAGTTAACGATA 1020
QY 977 CGCTGCGCGGAGAGGCTTTCTTGTGCTGTGTATGATGTAACCGACACTTTGG 1036
DB 1021 CAGTAAAGAGAACGGGATGCTTCCCGTGTGACAGTATATCCGCGCACATATGCG 1080
QY 1037 ATCAGATGACAGGATTTCTGCAACTGATGATGATGATGACGACAAAATCTATTGG 1096
DB 1081 ATCAGATGACAGGATTTCTGCAACTGATGATGATGATGATGATGATGATGATGAT 1140
QY 1097 TTGGGCTCAACCAAGGATTTGCTCAATGTTGAGAGCAGCAAAAGATTAACAAATGC 1156
DB 1141 TTGGGCTCAACCAAGGATTTGCTCAATGTTGAGAGCAGCAAAAGATTAACAAATGC 1200
QY 1157 AGAATCTATTAATACAGTGTGCTGCGGAGGCTTTTCCAGTGGGCGGCTGAATCTGTG 1216
DB 1201 AAAATTAATCTTGTGCAATATGATGATGATGATGATGATGATGATGATGATGATG 1260
QY 1217 CCAGCTTGAACGAGAAAGAACTAGAGGCTGCGGAGGCTGCTTATGATGATGATGATG 1276
DB 1261 AAGATCTTGAACGAGAAAGAACTAGAGGCTGCGGAGGCTGCTTATGATGATGATGATG 1320
QY 1277 GCTGGGCTTCAAGAGCCAGAAATCAATCATCTTCAAGAAAGCTGTGATCGCAACAA 1336
DB 1321 TGTGGGCTTCAAGAGCCAGAAATCAATCATCTTCAAGAAAGCTGTGATCGCAACAA 1380
QY 1337 TTAAAGAAAGTACTGCGCTTTGATCATTTTGTGATTTCCAGCCCTTACAGAGCCAG 1396
DB 1381 TCGTAAAGTCTGAGGCTTTTATGAGGCTTTCCAGTGTATGATGATGATGATGATGAT 1440
QY 1397 TCGATATGAGGCTTCCGCGTATGAGCTCAAGCTGCTGTTGAACCA---ACTGTCAACCCG 1453
DB 1441 TCGCATGTGCTGAGGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
QY 1454 CACCGGCTATTAACATGAGCCGATGTGAGATCTGCGTGTTCACAGCAAGAAAGCTGAAG 1513
DB 1501 TCGTCAAGTCTCCGAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
QY 1514 AAGTGTGAGAGGAGAAAGATCAAGAAAGCTTCCAGCCCTGCTCCCTGAAATGAAA 1573
DB 1561 AGAATCTCAAGGAGGAGAAAGCTCCAGAAAGCTTCCAGCCCTGCTCCCTGAAATGAAA 1620
QY 1574 AAGAGACCTGAGAGGAG-----AAGTACACCTCATTTATGAGAGGAGGAGAGCAG 1624
DB 1621 TCGAGGAGCTGCGAGAGTGTCTGCGAGATGAGAGGAGCTTCCAGGCGAGACCGGAGAG 1680
QY 1625 GTAGCGTGAAGACCAAGAGACATCAATGAGGATCAAGTTTACCAAGGAGAGAGAGA 1684
DB 1681 CACTCGTGAAGACCGCGGCTCATGTAAGATTAATCACTCAAGCAAAATGACCGATGTA 1740
QY 1685 TTGGGCTTACGCTATTAATTTCAACCCAGGCGTATTTGAATGATGAAAACTGGCGTGA 1744
DB 1741 TCGGACAGTATATGTTGTCTCGCCGATCTTGTGTGAAGAACGCTTAACTGCAACAG 1800
QY 1745 TCCACCATTTGGCGGAACAGTACTGATATGATCAAAAGTATGAGGAGGAGAGATACA 1804

```

1801 CACACCCCTGACACACAGGTTAAGATCATACCCACTCCGAGAGATCAGAAAGTATG 1860
1805 AAGTGAGCATTACACCGGTAGGTCATTGTACCGAAGAGGAGCGCGCTGTTCAAG 1864
1861 CAGTGAAACCATACACCGCTTAAGATCTGANTGCCAGAGAAAGTCCCGATCATGCGCAG 1920
1865 ACTTCAGGCAATTGAGTGAAGCGGTAGCATGTTTCAACGAGAGGAGTTCGTAACA 1924
1921 AATTCTTAGCACTGAGTGAAGCGCCAGCTTGTTGACACGAAAGAGATTGTGAAC 1980
1925 GATACCTGACCACTCGCATCAACGAGAGCGCCTAAACCTGACGAAAGTACTATA 1984
1981 GCAACCTGTACCATTTGTCATGACACGCTCCGCTAAGAAATACAAAGAGGAGCGTACA 2040
1985 AGACTGTAAAGACTCAGGACACAGACTGAGATAGCTCTGATTTGACGACGAAAGT 2044
2041 AGGTTACAAAGGCAAGCTCCGAGAAACAGAGTACGTTGTGACGTGACAAAGACGAT 2100
2045 GTGTTAAGCGAAGACGACAGGTCCCTTGCTTAACCGGTGATCTGTAATCCACAT 2104
2101 GCGTTAAGAGGAAAGAGCTTAGAGACTTGCTTTCGGAAGACTGACCAACCCGCT 2160
2105 TTCAGAGTTTGCCTGACGAGACTCTCAAGACGACGACGACCTGACAAAGTCCAA 2164
2161 ATCAAGAACTAGCTTGTAGAGGACTGAAGACTGACCCGCGGTCCGTAACAAAGTTGAAA 2220
2165 CCATCGGAGCTGATGAGATGCGACAGTTCAAGTAAATCTGGAATCATCAAAAGCGTGTGA 2224
2221 CAATAGGAGTATAGGACACACAGAGATCGGGCAATCAGTATCATCAAGTCACTGTCA 2280
2225 CTAAAGAAAGATCTGTTGTGAGTGCAGAAAGAAAGAACTGCGCAGAAATCATCAGGATG 2284
2281 CGGCACTGATCTTGTACACGCGGAAAGAAAGAAAGTCCGCGCAATTTAGAGCGCGACG 2340
2285 TAAAGAGATGAGACGATGATGATGTTGCTGCTAGGACTGTGATTCAGTCTTCTAAATG 2344
2341 TGCTACGGCTGAGGGGACATGACATCAGTCAAGTCAAGTGAATTCGGTTATGCTCAACG 2400
2345 GGGTTAACCACCCCTTAACTCTGTACATTTGATGAGGCAATTCCTGCTGCTGACGGA 2404
2401 GATGCCAAGAACCCCTGAGAGTGTGATGTTGAAGAAAGCGTTCCGATGCCACGAGAG 2460
2405 CGCTGCTGGCACTGATGTCATGTCATCAAC---TAAAGAAAGTGTATTGTGCGGGAGC 2461
2461 CACTACTTGCTGATTTGCAATGTCACAGCCCGTAAGAGTACTATGTCGAGAGC 2520
2462 CAAAACAAATGCGGCTTCTTAAATGATGTCCTGAAAGTACATTTTAACT----- 2512
2521 CTAAGCAATGCGGATTTCTTCAACATGATGCAACTAAAGTACATTTCAACCACTGAAA 2580
2513 ATGACATATGCACTGAATGTATCCATTAAGAGCATCTTACGAGGTGACACAGACTGTA 2572
2581 AAGACATATGTAACCAAGCATTTCTCAAGTTATCTCCGACGTGACACAGCGACATCA 2640
2573 CCGCATGCTGCTCCACGCTCTTCTAGCAACAGCGAATGAAGCGGTTAAACCATGCTG 2632
2641 CCGCATATGATTCACACTGCTATTCGATGGAAGAAATGAACCAAAACCCGTCAGAA 2700
2633 AATAAATCATGATAGATACCAACAGGACACAAAGCCGACAAAGATGATCTGATTTCA 2692
2701 AGAACATGAAATGACATTTACAGGGGCGACGAAACCGAAGCAGGGGACATCATCTCGA 2760
2693 CCTGTTTAAAGAGATGGGTGAACAGCTACAGATTGACTACAAAAATACAGAAATCATGA 2752
2761 CATGTTTCCGCGGGGTGTTAAGCAATGCAATGACTATCCCGAGATGAGGTAAAGA 2820
2753 CTGCGGCTGATCGCAAGAACTTACGCGGAAAGCGGTTATGCTGCTGAGGTAAAGTCA 2812
2821 CAGCGCGGCTTCAAGAGGCTTACCAAGAAAGAGTATGCGCTCGGCAAAAGTCA 2880
2813 ACGAATCACTACTGACGACTTGTGACAGTGAACGTGTACTTACACGACAG 2872

2881 ATGAAGCCCGCTGTACCGCATCATCAGAGCATGTGAACGTGTGCTCACCCGACATG 2940
2873 AAAAAGCATTTGTGGAAGACGTAAGTGTGATCCCTGATTAAGACATTTACAGCTA 2932
2941 AGACAGGCTATGTAAGAACTTTACAGGGCGACCAATGATTAAGACGCTCAATAAG 3000
2933 AATATCCGGGAGATTTACAGGCTTCAATGGAACGATGCGACGCGGACACAGACGCAATTA 2992
3001 TACCTAAAGAAATTTTCAAGGCGACCATCGAGATCTGGAAAGCTGAAACAAAGGAATTA 3060
2993 TGGACGCGTCTTGTATAGCCCGACAGACGCTGATGTGTTCCAGAAATAGGTGAACGCT 3052
3061 TTGCTGCATTAAGACGTCCTCCCTCCCTACCAATCCCTGACGTGCAAGACTTAAGCTTT 3120
3053 GCTGGCGAAGGCTTTAGAGCAGCTTGGCGACGCGCAACATTTGCTGACGACAGC 3112
3121 CTTGGCGAAGGACACTGAAACGATCTGGGCGAGCGCGGTATCTTACTTACCGGTTGCC 3180
3113 AGTG---GAGACGTTGACCCCATTTCAAGCATGACAGAGCGTACTCACCTGAAATGGCAC 3169
3181 AGTGAGGAGAGCTGTTCACAGTTTGGGATGACAAACACACTCGGCAATCTAGCGCT 3240
3170 TGAATCTTTTGCACAGGTTCTTTGGATGAGACTGGAAGTGGTTATTTTCGCTC 3229
3241 TAGAGTATTTGCAATTAAGTTTTCGCAATGACTTGAACAAGCGGCTGTTTCCAAAC 3300
3230 CTACCGTGCATTAATCTTACAGGAT-----CAGCACTGGGATA 3268
3301 AGAGATTCCTGTTAAGTACATCATCTGCGCATCAGCGAGCCAGTACCTATGGGACA 3360
3269 ACTGCGCAGGAAACATGATGAGGCTTAATAGAGAGTGAACAAAGAGTTGTACAGCG 3328
3361 ACAGCCAGAAACACAGAAATATGGATGATGATCAGCCGTTGCCGCAACTCTCCCTA 3420
3329 GATATCGGTGATCAAAAGCGGTTGACAGAGCGAGTGAAGTATTAAGAAATATA 3388
3421 GATTTCCGCTGTTCCAGCTAC---TGGAAGAGCACACACTGATTTGACAGAGCGCA 3477
3389 CCATCAAGGACTACTCTCCACAAATTAATGTGTTCCATTAAATGCGCGGTGCCCCCT 3448
3478 GAACTAGATTAATCTGTCACAGATTAATCTTGTCCTGATACCGCAATCTCCCTCACG 3537
3449 CGTTGATGTTGACACAAAGAGAGGATCAACTGATCAGCGGATTTCTATCTAAGA 3508
3538 CTTATGTCCTCGAGACACAGAGAAACACCCGCGCGGTGAAAATTTCTTGAAGCCAGT 3597
3509 TGAAGGCAATCTGTGTGTGATCGGCACTCTA-----TCAGCATTCAGAGG 3558
3588 TCAAAACCACTCTCGTACTTGTGATTCAGAGAAATAATGAAGCTCCCAACAAAGAA 3657
3559 AAGAAATGAGTGCATGCGGTCCATTTGCCACTAATACATCAGGTGATCTCGATTTG 3618
3658 TCGAATGATGCGCCGATTTGAGATGCGCGGAGATTAACATCAACCTGCGTTTCG 3717
3619 GGAATACCTAGCACTGTCGTAATATGACATTAATCTTGTCAATGTTAAGAGCCCGTAC 3678
3718 GGTTCGCGCGCAGACGCG---TACGACTGTGTTTCAATATTTGAACTTAATATAC 3773
3679 AGAACAATCACTACCAACAGTGCAGAGATCAGCTATCCACACAGCATCTACGCTG 3738
3774 AGAAACATCACTTTCAACAGTGCAGAGACACCGCGGACCTTGAATAAACCTTTCCGCT 3833
3739 AAGGCTGTCCACACTGAAACACTGCGGAAACATGTGTGCTATAGGGTATGGGCTTCT 3798
3834 TCGGCGCTGAATCTCTTAACCCCGAGGCACTCTGCGTGAAGTCTCTAGGTTACGCC 3893
3799 GATCGCGCAACGAGAAATATCATCATGCGGTGACGCTCATTTAGGTTAACCGGTC 3858
3894 GACCGCAATATGAGAGAGTGTACCGCTTTTCCGAAATTTTGTCAAGGTGTCTCA 3953
3859 TGTACGCTTAAGAACACTGCGGAATATCTGAGGTTCTTGTGTTCTTCCGCAAGGAC 3918
3954 GCGAGGCGAGAGTGTCTCAAGCAATACAGAAATGATCTGATTTTCCGACAACTAGAC 4013

QY 323 TTAAGAAA-----GTGACATTACCGACAGAAATAGCTCTTAAGCGCGAGACTGEC 376
Db 364 TGGCGAAAAGACATGTAAATTTACAAACAGAACTTGATGAGAAAGATCAAGAGCTCC 423
QY 377 TGGAAATCATGTCAACACGACGACGAGACTTCATCTTGATGACACAGAGCCCA 436
Db 424 GGACCGTACTTGATACACCGGATGTGAAACCGCATCTGCTGCCCAACGATGTTA 483
QY 437 CGTGTAGGTACTTTGGAAGTGTAGAGTATACCAAGATGTAGCGAGTCCATGACGGA 496
Db 484 CCGACGACACGCGTCCGAGTACTCCGTATAGCAGAGCGGTGACA--TCAACGCTCCCG 540
QY 497 CATCAATCTACCAACGACGCGCTTAAAGAGTTAGACATTTACTGATAGGCTTTGACA 556
Db 541 GAACATTTTACCAACGAGGTATGAAAGGCGTGGACCCCTGTACTGATGCTTGCAC 600
QY 557 GCAGCCCTTTTATGTACAAAACATGCGAGGTTCTTACCCTTACTTAAACAGAACTGGG 616
Db 601 CCACCCAGTTCATGTTCTGCGCTATGGCAGGTTCCGATCCCTGATTAACCAACCACTGGG 660
QY 617 CTGAGAGAGAGTATTGGAAGCAGTAACATGAGCTGCTGCTGTAACATGATCTTCAGAGA 676
Db 661 CCGAGAAAAGAGTCCCTGAAAGCGTAAACATGCGACTTGACAGCAACAAAGCTGAGTGAAG 720
QY 677 GCAGGCTTGAAGAACTCTCAATCCTTAGAGAAAGAGGCTCCAACTTACTAATAAGATCA 736
Db 721 GCAGAGACAGAAAGTGTGATATATGAGAAAGAAAGATGAAAGCCGGGTACAGGGTTT 780
QY 737 TATTTCCGTTGTTTCAACATCTACAGAAAGATAGATCACTGTTACGTAGCTGGCATC 796
Db 781 ATTTCCTCGTTGATGACACTTTTACCAAGAACAGAGCCAGTTGACAGACTGGCATC 840
QY 797 TTCCAAGGTGTTCACTTTGAAAGAAAGTCTAATCTACAGGTAGATGTGGACCATTTG 856
Db 841 TTCAATCGGTGTTTCACTTTGAAAGAAAGCACTGTAACCTTGCCTGTGTATTAAGTGG 900
QY 857 TCAGCTGTGAAGGCTACGTCATCAAAAGATACATCAGCCGAGACTATACGTTAAAG 916
Db 901 TGAGCTGGGAAGGCTACGTAAGTGAAGAAATCAACATCAGTCCGGGATCAAGGAGAAA 960
QY 917 TTGAGACTTGGCGTCCACATGATCATCGGAGGGTTTCTTGATGTTCAAAAGTCAAGATA 976
Db 961 CCGTGGGATACGCGGTTTCAAAACATGAGGAGGGCTTGTGCTATGCAAAAGTTACGATA 1020
QY 977 CGCTGGCGGCGAGAGGTTTCTTTGCTGTGTATGATATACAGCCACACTTTGGC 1036
Db 1021 CAGTAAAGGAGAAAGGGGTATCGTTCGCCGTGTGACGTATATCCCGGCAACCATATGCG 1080
QY 1037 ATCAGATGACAGGAGTTCTGGCACTGACGTTAGTGTGATGACGACAAAACATATTGG 1096
Db 1081 ATCAGATATCCGCAATTAATGCGACGAGATATCTACCTGACGATCACAAAACCTTTGG 1140
QY 1097 TTGGGCTCAACCAAGAGATTGTCTCAATGTGTAGACGCAAGAAATATCTAACACATGC 1156
Db 1141 TTGGGCTCAACCAAGAGATCGTCAATTAACGTAAACATCAACGAAACCAATATACATGC 1200
QY 1157 AGAATCTATATTAACAGTGTGTGCGCCAGGCGTTTCCAGGTGGCGGCTGAACATCGTG 1216
Db 1201 AAAATTTACTCTTGCAATTCATTTGCACAAGGTTTACGAAATATGGGCGCAAGGCCAAG 1260
QY 1217 CCGACTTGAAGCAAGAAAGAACTAGGAGGTGCGGAGGCGCACTTTACTATAGGCTGCT 1276
Db 1261 AAGATCTTGACATTAATAAAATGCTGGGACCAAGAGAGGCAAGCTTTACATATGCTGCT 1320
QY 1277 GCTGGCTTTTAAAGACCAAGAAATCACATCATCTACAGAAAGCTGTAGCGCAAGAA 1336
Db 1321 TGTGGCGCTTTGCGACTAAGAAAGTGCATGTTTATGCGCCCACTGGAACGAGACCA 1380
QY 1337 TTAAGAAAGTACCTGCTTTGATCTATTTGTATTTCCAGGCTTACAGCCAGGAGC 1396
Db 1381 TGTAAAGATCCCAAGCTCTTTTATGCGCTTCCCATGTATCCGATGATGATCTTCTT 1440

QY 1397 TCGATATGAGGCTTCGCCGTAAGGCTCAAGCTGTCTGTGAACA---ACTGTCAAAACCG 1453
Db 1441 TGCCCATGTCCGTGAGGAGAAAGATGAATTTGGCATTTCAACCAAGAAAGAGGAAAAAC 1500
QY 1454 CACCGCTATTTAATGATGCGATGTGAGCATCTGCGTGGCTTTACAGCAAGAGCTGAAG 1513
Db 1501 TGTCTCAAGTCCCGAGGAAATTAATGATGAGGCCAAGGCTGCTTTCCAGGATGTCTCAG 1560
QY 1514 AAGGTGTCGACGGGAAGATCAAGAGAGCCCTGCCACCTTGCTCCCTGAATATAGAA 1573
Db 1561 AGGAATTCAGAGCGGAGAAAGCTCGAGAGAGCACTCCACCAATTAGTGCAGCAAGGTA 1620
QY 1574 AAGAACCGTAGAGCAG-----AAGTAGACTCTATTATGCAAGAGCGAGAGCAG 1624
Db 1621 TCGAGGACGCTGCGGAATTTGTCTGCAAGTGTGAGAGGGCTCCAGCGGACACCGGAGCAG 1680
QY 1625 GTAGCTGAGAGACACCAAGAGACATCAGGGTGAACAAGTTTACCAGCGAAGAAAGA 1684
Db 1681 CACTGTGGAACCCCGCGGCTCATGTAAAGATTAATCTCAACCAATGACCGTATGA 1740
QY 1685 TTGGGTCTTAAGCTACTTTACCCCGAGGGGTATTGAATATGTAATAAACTGGGCTGA 1744
Db 1741 TCGACAGTATATGTTGTCTGCGCGATCTGTGTGAAGAACGTTAACTGCGACAG 1800
QY 1745 TCCACCCATTTGGCGGAAACAATGTAATGTAATGACTCAAAAGTGTAGGCGAGAGATACA 1804
Db 1801 CACACCCGCTAGCAGACCAAGTTAAGATCATACCACTCCGGAAGATCAGGAAGTATG 1860
QY 1805 AAGTGAGCCATPACCAAGGTAAAGTCAATTGTACAGAAAGGACGCGGTTCCCTGTTCAAG 1864
Db 1861 CAGTGCAGACATPACGACCTTAAAGTATGATGCGACAGAAAGTGCCTTACATGCGCAG 1920
QY 1865 ACTTCCAGGCAATGATGATGAGCGCTACGATCGTTTCAAGAGAGGAGTTCTGAACA 1924
Db 1921 AATTTTACACTGATGATGAGCGCACGCTTGTGTAAAGAAAGAAAGATTTTGAACC 1980
QY 1925 GATACCTGCAACCATCGCAATCAACGAGAGGCGCTTAAACATGACCAAGATCTATTA 1984
Db 1981 GCAAGCTGTATCATATTTGCCATGACAGGTCCCGCTAAGAAATACAGAAAGAGCATGCA 2040
QY 1985 AGACTGTAAAGACTCAGAGACAGATCTCAAAATACGTTTGTGATATTTGACGACGAAAGT 2044
Db 2041 AGGTTACAAAGCAGAGCTCGCAAGAACAGAGTACGTGTTTGAAGTGAACAAAGCGAT 2100
QY 2045 GTGTTAAGCGAAGAGCGCAGAGTCCCTGTGTCTTACCGGTGATGTGTATGATCAACAT 2104
Db 2101 GCGTTAAGAAAGAAAGACCTCAGAGACTTGTCTTTCCGAGAAACTGACCAACCGCCCT 2160
QY 2105 TTCAAGGTTTGGCGTAGAGAGTCTCAAGACAGCAGCAGCACTCACAAGTCCGA 2164
Db 2161 ATCAGAACTTACTTGTAGGGAAGTGAAGACTTCAACCCGCGTCCGTAAGAGTTGAAA 2220
QY 2165 CCATCGAGTCTATGAGATGCGCAGGTTCAAGTAAATCTGAATCATCAAAAGCGCTGTA 2224
Db 2221 CAATAGAGATGATGAGCACAACAGGATGTGGCAAGTCAATCATCAAGTCACTGTCA 2280
QY 2225 CTAAGAAAGATCTGTTGTGTGAGTGAAGAAAGAAATCTGCGCAAAATCATCAGGGATG 2284
Db 2281 CCGACAGTATCTTTGTTACCGCGGAAAGAAAGAAATCTGCGGAAATTTAGAGCGGAGC 2340
QY 2285 TAAAGAGATGAGAGTATGATGTGTCTGTAGAGCTGTGATTCAGTGTCTTCAATG 2344
Db 2341 TGTATCGCTAAGGGGATGTGATACGTGTGAAGACAGTGAATTCGTTATGTCTAAGC 2400
QY 2345 GGGTTAAGCACCCGTTTAAACCTGTATCATTTGATGAGCAATTTGCTGCCATCAGGGA 2404
Db 2401 GATGCACAAAAGCGGTAAAGTGTGTATGTTGAAGAAAGCTTCCGGTGCAACGAGAG 2460
QY 2405 CGCTGTGAGCACTGATTTGCTGTCAAC--TAAAGAAAGTGTATTTGCGGGAGCC 2461
Db 2461 CACTACTTGCCTTATATGCAATGTCAAGCCCGTAAAGAGTGTACTATGCGGAGACC 2520
QY 2462 CAAAAACATGCGGCTTTTAAACATGATGTGCTGAAGATCATTTTAAAC----- 2512

```
Db 2521 CTAAGCATGGGATTTCTCAACATGATGCAATTAAGATCACTTCAACCAACCTGAAA 2580
Qy 2513 ATGACATATGCACTGAAGTGTACCAATTAAGACATCTCTAGAGGTGCACACAGCTGTA 2572
Db 2581 AAGACATATGATACCAAGACATTTCTACAGTTATCTCCGAGTTTGCAACAGCCAGTCA 2640
Qy 2573 CCGCATATGCTCCAGCTCTTCTACGCAAGGAAATGAAGACGGTTAAACCATGTGTG 2632
Db 2641 CGGCTATTTGTATCGACACTGATTAAGATGAAAAATGAACCAACAAACCCGTGCAAGA 2700
Qy 2633 ATAAATATCATATGATATCAACAGAGGACCAACAAACCCGACAAAGATGATCTGATTTCAA 2692
Db 2701 AGAATATGAAATGACATTTACAGAGGGGCAAGAAAGCCGAACCCAGGGGACATCATCTGCA 2760
Qy 2693 CTTGTTTACAGAGATGGGTGAACAAGTCAAGATTTGACTACAAATATCAAGAAATCATGA 2752
Db 2761 CATGTTTCCGGGTGGGTGAAGCACTGCAAAATCGACTATCCGGAATGAGGTAAATGA 2820
Qy 2753 CTGGGCTGCATTCGCAAGGACTTAACGGGAAAGGCTTATGCTGACAGTACAAAGTCA 2812
Db 2821 CAGCGGCGGCTCAACAGGCTTAACAGAAAAAGATGATGCGTCCGGCAAAAAGTCA 2880
Qy 2813 ACAGAAATCCACTACTCGACACTTCTGAGCAGTGAACGTGTACTTACAGSCACAG 2872
Db 2881 ATGAAAACCGCTGATACCGCATCATACAGACATGTAAGTGTGCTCAACCGCACTG 2940
Qy 2873 AAAAAACGATTTGTGGAAGACGCTAGCTGTGATCCCTGATTAAGCACTTACAGTGA 2932
Db 2941 AGGACAGGCTGATGTAAGAACTTTCACAGGGGCAACCCATGATTAAGCACTTACAGTGA 3000
Qy 2933 AATATCCGGGGGATTTTCAACGGCTTATGAGCAGCTGGCAGGGGAAACAGACGCCATTA 2992
Db 3001 TACCTAAAGGAAATTTTCAGGCACTGAGAGATGGGAACTGAAACAAAGGAATTA 3060
Qy 2993 TGGCAGCGGTTCTTGATTAAGCCGAGACAGCTGATGTGTTCCAGAAATAGGTGAACGTCT 3052
Db 3061 TTGTGGGATTAACAGTCCGCTCCCGTACCAATCCGTTCAAGTCGCAAGACTAAGTTT 3120
Qy 3053 GCTGGCGAAGGCTTTAGAGCAAGTCTTGGCCACGGCCAAACATGTGCTGACAGACAGC 3112
Db 3121 GCTGGCGAAGGCACTGGAACCGGATCTGAGCCAGCGCGGTATGCTTACCTTACCGTTGCC 3180
Qy 3113 AGTGG---GAGAGCTGACACCATTCAGACATGACAGAGCGTAACCTCACTGAATAGGAC 3169
Db 3181 AGTGAAGGAGCTGTTCCACAGTTTGGGATGAAACCAACACTCGGCCATCTACAGCTCT 3240
Qy 3170 TGAATCTTTTGGACCAAGTTCCTTTGGAGTGAAGCTGGACAGTGGGTTATTTTCCGCTC 3229
Db 3241 TAGAGGTAAATTTGCACTTAAGTTTTCGGCATGAGCTTGAACAAGCGGCTGTTTCCAAAC 3300
Qy 3230 CTACCGTGCACCTTACTTACAGGAT-----CAGCATGGGATA 3268
Db 3301 AGACATATCCGTTAAAGTACATCTGCGCATGCGAGGACAGTACCTGATGGGACA 3360
Qy 3269 ACTGGCAGGGAAGAACTGTATGGGCTTAATAGAGAGTAGCAAGAGATTTGCACGGC 3328
Db 3361 ACAGCCGAGAAACGCAAGTATGGGTACGATACAGCGGTGCGCCGAACCTTCCTGTA 3420
Qy 3329 GATATCCGTGATCAAAAGCGGTTGACAGAGGAGGTAGCTGATTAAGAAATATA 3388
Db 3421 GATTTCCGCTGTCCAGCTAGC---TGGGAAAGGACACAGCTTGAATTTGCAAGCGGCA 3477
Qy 3389 CCATCAAGACTACTCTCCAAACATTAATGTGTTCCATTAATGCCGGTGGCCCACT 3448
Db 3478 GAACTAGAGTATCTCTGACAGCAATTAAGTGTCCAGTGAACCGCAATCTCCCTCAG 3537
Qy 3449 CGTTGATGTTGACCAAGAGACAGGGTGAACAGTATCAAGCGGATTTCCATATAGA 3508
Db 3538 CTTAGTCCCGAGACAAAGAGAAACAACCGGCGCGGTGAAAAATTTTGAAGCCAGT 3597
Qy 3509 TGAAGGCAAAATCTGTGTTGGATCGGCGATCTTA-----TCAGCATTCAGAGG 3558
Db 3597 TGAAGGCAAAATCTGTGTTGGATCGGCGATCTTA-----TCAGCATTCAGAGG 3558

Db 3598 TCAAAACCACTCCGTACTTGTGATCTCAGAGAAAAAATTGAAGCTCCCAAGAGAA 3657
Qy 3559 AAGAAATAGATGATCCATGGGTCCATGTCGCCATTAATACATCAGGTGATCTGATTTG 3618
Db 3658 TCGATGATATGCGCCGATTTGGATAGCCGGCGAGTAAAGACTACACACTGGCTTTG 3717
Qy 3619 GGAATACCTTACGATGTGGTAAATATGACATTAATCTTTGTCAATGTAGAGCCCGTAC 3678
Db 3718 GGTTCGGCGGACGGACCG---TACGACTGGTGTATCAATATGGAATTAATATAC 3773
Qy 3679 AGGAACCATCACTTACCAACGTGGAGAGATCAAGCTATCCACACAGATCTTAAGTGT 3738
Db 3774 AGAAACCATCACTTCAACAGTGGAGAGACAGCGCGCACTTGAAGAACCTTTCGGCT 3833
Qy 3739 AAGCTGTCCACCACTGGAACACTGGCGGAAACATGTGTGCTTAAAGGTATGGGCTTGTCT 3798
Db 3834 TCGGCTCTGAATCTGCTTAACCCCGGAGGACCTCTGTGTGAATGCTTACGTTACGCC 3893
Qy 3799 GATCGGCAACCGGAATATCATCTGCGGTGGCAGCTCATTTAGTTTACCCGCTGC 3858
Db 3894 GACCGCAATATGATGAGACGTATGTCACCGCTCTTGGCAAGAAATTTGTCAAGTGTGCA 3953
Qy 3859 TGTACGCTTAAGAACTCTGCCAAATATCTGAGTTCTTTCGTTCTTCCGCAAGAC 3918
Db 3954 GCGAGGCAAGAGTGCCTCAAGCAATACAGAAATGATCTGATTTCCGACAACTAGAC 4013
Qy 3919 AACGGCAACCAACACATGACAGAGACAGACTGGGTATGCTTACCAATCTATCAA 3978
Db 4014 AACGCGGCAACAGCAATTTACCCGCAATTTGAATGTGTGATTTGTCCTGCTTAC 4073
Qy 3979 GGGTCAACAGATGACAGAGGAGAGCTTCAAGCTGACAGATGATCAGAGTGAATTT 4038
Db 4074 GAGGATCAAGAGACGAGATTTGAGCGGCAACCGTCTGATCTTAAGAGGAAACATT 4133
Qy 4039 AGCAAGCGCTGACCAAGCTATCTGTTAATGTCTGCTAATAGCAAGTCAACAGGTTCC 4098
Db 4134 GCTGATTTGAAGAGGAGAGGATGTCATATGACGCCATCACTGGGCAAGAGGAA 4193
Qy 4099 GGAATGTGGCGTGCAGTATCCGAAATAGCGCGGCTTTTGAATACAGCCAAATAGCT 4158
Db 4194 GGAATGTGGCGTGCAGTATCTTAACGTTGGCCGAACAGTTTACCGATTCAGCCACAGAG 4253
Qy 4159 GTCGGACCGGCTAGACT---TGTGAAGCAAGAACCGCTCATATCATGCTGTAGAGCC 4215
Db 4254 ACAGGTACCGGAAACGACTGTGTGCCAAGAAAGAGATGATCCAGCGGTTGGCCT 4313
Qy 4216 AATTTTCTAAGATCCGGAACCGGAGGCGAATTAAGTCCGAGCTGCTTACATGAC 4275
Db 4314 GATTTCCGAAACACCCAGAGGAGAGGAGCCCTGAATTTGCTGCAAAACGCTTACATGCA 4373
Qy 4276 ATAGCGTCAATCGTCAACGCTGAGCGGATTAACAAAATATCAAGTACGCTACTGTCAAC 4335
Db 4374 GTGGCAGACTTAAGTAAATGAACATTAATCAAGTCTGTGCGCATCCCATGCTATCTTACA 4433
Qy 4336 GGCATCTATTTCTGGTGGCAAAAGATCGAGTATGCAATTTGATCACTGTTTCACTGCT 4395
Db 4434 GGCATTTACGAGCGGGAAGAACCGCTTGAAGTATCACTTAATCTGTTGACAAACGGCG 4493
Qy 4396 TTGACACTACGATGCGGATGTCACCATATATTTGCTTGAATAAACATGGAGACAGG 4455
Db 4494 CTAGACAGAACTGATGTGGAGAGTAAACATCTAGCTCTGGAATGAAGAGTGAAGAAAGA 4553
Qy 4456 ATATACGAGCGCATTCACCCGAAAGAAAGGTC---GAAATTCGATGATGACAAAGCA 4512
Db 4554 ATGACACGGGTGCTCAACTTAAGAGATCTGTAACTGAGCTGAAGATGAGATATGAGAG 4613
Qy 4513 GTAGACATTTGACTTGTCAAGGCTCAACCAACAGCTCTTTGGCAGGACAGACAGCTTAC 4572
Db 4614 ATGACACAGAGTATGATGATTCATCCGAGACGTTGCTGAAAGGAAAGAGGATTC 4673
Qy 4573 TCGGTCAATGAGGCAAGTGTATTCATACCTGGAAGTACAGATTCATCAGACGCC 4632
Db 4674 AGTACTCAAAAAGAAAGTGTATTCGTACTTTGAAGGCAACAAATTCATCAAGAGCA 4733
```

QY 4633 AAGACATTCGCGAATCCATGATGCGCCCAAAATCTGAGGCTAATGAGCAGATT 4692
DB 4734 AAGATATGCGGAGATAAAGTCTCTGTTCCAAATGACAGAGAACAGCAACTG 4793
QY 4693 TCGTTTATCATCTCGGGGAGAGATGTCACAGATCCGCTCCAAATGCGCAGTAGAGAG 4752
DB 4794 TGTGCTCATATGATGGGGAGAGCATGGAAGCATCCGGAATAATGCGCGGTGACAC 4853
QY 4753 TCAGAGGCTGCTGCTCCACCTTCACACTTCCATGCTGTGTATTAAGCTATGCGCT 4812
DB 4854 AACCGTGTCTAGCCCGCCCAAAAACGCTGCGCTCTGTATGATGACATGACGCA 4913
QY 4813 GAGCGCTATACAGGTTGCGCTCTGCGAAGAAAGACGTTGCGGTAATGCTCATCTTC 4872
DB 4914 GAAAGGTCACAGACTCAGAGCAATACGTCAGAAAGATGACATGCTCTCCAC 4973
QY 4873 CTGTTCCGAAATACAGATCAAGCGCTGCAAGATACAGTACAGCAAAACAGTCTG 4932
DB 4974 CCCCTTCCAAAGTACAAATCAAGATGTTGAGAGGTTGAGTGCACAAAGTAGTCTG 5033
QY 4933 TTTTCAGCGCTGCTACACCGGCTTACACCCGAGAAAGTACGCGAA 4980
DB 5034 TTTAACCGCATACCCCGCATTCGTTCCCGCGGTAAGTACATGAA 5081

RESULT 10

US-08-801-263A-7

Sequence 7, Application US/08801263A
Patent No. 5811407
GENERAL INFORMATION:
APPLICANT: Johnston, Robert E.
APPLICANT: Davis, Nancy L.
APPLICANT: Simpson, Dennis A.
TITLE OF INVENTION: System for the In Vivo Delivery and
NUMBER OF SEQUENCES: 12
TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
STREET: 1211 East Morehead Street
CITY: Charlotte
STATE: No. 5811407th Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,263A
FILING DATE: 19-FEB-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-147
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 11663 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-801-263A-7

Query Match 26.7%; Score 1333.2; DB 1; Length 11663;
Best Local Similarity 56.1%; Pred. No. 0;
Matches 2820; Conservative 0; Mismatches 2128; Indels 80; Gaps 13;

QY 23 ATATGAAAGAAATTCACGTTGACTTAGTGTGACAGCCCGTATGTCAAGTCTTACAGC 82
DB 64 AAGAGCCAGTATGTTACCTAGACGTAGACGCTCAGAGTCCGTTTGTCTGTGCAACTGCAA 123
QY 83 GAGCGTTTCCAAATTTGAGATCGAAGCAGAGCTCACTGACATGACCAATGCAATG 142
DB 124 AGAGCTTCCCGCAATTTGAGTAGTACAGACAGAGCTCACTCAAAATGACCATGCTAATG 183
QY 143 CCAAGCGTTTTCGATGTGCAACAAAGCTCATTTGAAAGGAAAGTCCACCGGAGCAAG 202
DB 184 CCAGAGCAATTTGATGTGCGACAGTAACTGATGAGAGTGGAGTTCTTACACAGGCA 243
QY 203 TTATCTGGAATTTGAAAGTGGCGCCGTCAGACATGACATTTCCATTCACCGCTATCAT 262
DB 244 CGATTTTGGACATGAGCAGCGCAGCGCTCGTAGAATGTTTCCAGACATGATACAT 303
QY 263 GTATCTGCGCTATGATAGCTGAGAGCCGCGACAGACTGACATCAACAGTATGCAAGAG 322
DB 304 GCGTTTGGCCCATGCTGATGTCAGAGAGCCGAGCCGATGATGAAATATGCGACAAAC 363
QY 323 TTAAGAAA-----GTGACATTAACCGACAGAACTAGCTCTTAAGCGGAGACCTGC 376
DB 364 TGGCGAAAAAGCATGTAAAGATTCAAAAGAACTTGATGAGAAAGATCAAGAGCTTC 423
QY 377 TGAAGTATGATCAACACAGCAGAGACTCCATCTGTGTATGACACAGAGCA 436
DB 424 GAGCGCTACTGATGATACCGAGTGTGAAAGCCATCACTGTCTTCCACACAGATGTTA 483
QY 437 CGTGTAGTACTTTGGAAGTGTAGAGATGATACCAAGATGTAGCAGCAGTCCATGACCGA 496
DB 484 COTGCAACAGCGTGGCGAGTATCCGTCATGAGAGAGTGTGAC--TCAAGCTCCG 540
QY 497 CATCATTTACACAGCGCTTAAAGAGTGAAGCAATTTATGATGATGAGCTTTGACA 556
DB 541 GAACATTTTACACAGCGCTTAAAGAGCGCGAGCCCTGTACTGATGATGCTTGCACA 600
QY 557 CGACCCCTTTTATGTACAAAGAAACATGCGAGGTTCTTACTTACTTAAACAGAACTGG 616
DB 601 CCACCCAGTATGTTCTCGGCTATGAGAGTTGTTACCTGATACAAACCACTGG 660
QY 617 CTGACGAGAGATTTGGAACACAGTAACTTTGCTGTGATCACTCAGATCTTTCAGAGA 676
DB 661 CCGAGAAAAGTCTTGTGAACGCGGTATCATCTGACGACCAAAAGCTGAGTGAAG 720
QY 677 GAGGCTTGAAGAACTCAATCTTGAAGAAAGAGCTCAACTTAATAAGATCA 736
DB 721 GAGAGACAGAAAGTGTGATTAAGAGAAAGAGATTAAGCCCGGTACCGGTTT 780
QY 737 TATCTGCTGTTGTTCAACATCTTACACAGAAATGATCACTGTATGATGCTGGCATC 796
DB 781 ATTCTCTCGTTGATCGACACTTTTACCCAGAACAGAGCCAGCTTGACAGAGCTGGCATC 840
QY 797 TTCCAAAGCTGTTCCACTTGAAGAAAGTAACTTCAAGTATGATGAGGAGCATTTG 856
DB 841 TTCCATGCTGTTTCCACTTGAAGAAAGAGTGTATCACTTCCGCTGTATACAGTGG 900
QY 857 TCAGCTGTGAAGGAGTGTATCAAAAGATGATGATGAGCCAGGACTTATAGGTAAG 916
DB 901 TGAGCTGTGAAGGCTAGCTAGTGAAGAAATCACCATGATCCGCGGATTCACGGAGAAA 960
QY 917 TTGAAGACTTGGCGTCCCAATGATGCGAGGTTCTTGAAGTGAAGTTCACAGATA 976
DB 961 CCGTGGATACCGGTTTCAAAACATGACGAGGCTTTTCTATGCAAGTTACCGATA 1020
QY 977 CGTGGCGGCGAGAGGTTTCTTTGTGTGTATGATGATGACAGCCACACTTTGCG 1036
DB 1021 CAGTAAAGAGAGAGAGGATGTTTCCCGGTGTGATATATCCCGGACCATATGCG 1080
QY 1037 ATCAATGACAGGATTTGCGAACTGACGTTAGTGTGATGACGACAAAATCTATTGG 1096
DB 1081 ATCAATGACAGGATTTGCGAACTGACGTTAGTGTGATGACGACAAAATCTATTGG 1140
QY 1097 TTGGGCTCAACCAAGATGTTGTCTCAATGATGAGGCAAAAGAAATCTAAACAAATGC 1156

```

Db      1141 TTGGGCTCAACGAGCAATCGTCATTAACGGTAAGATTAACGAGAACCCAAATTCATGTC 1200
Qy      1157 AGAACTATCTATTACCAAGTGTGCGCCAGGCGTTTTCAGGTGGCGCGTGAACATGCTG 1216
Db      1201 AAAATTATCTTCTGCGCAATATGACAGAGGTTTCAGCAATGGGCCAAGAGGCAAG 1260
Qy      1217 CCGACTTGAGCAGACGAGAAAGAACTAGAGGGGTGCGGAGCGCACTTACTATAGGCTGCT 1276
Db      1261 AAGATCTTGACATGAAAAATGCTGGGACACAGAGAGCGCAAGCTTACATATGCTGCT 1320
Qy      1277 GCTGGGCTTTCAAGACCCAGAAAATCAATCATCTACAGAAAGCTGTGACGCAACAA 1336
Db      1321 TGTGGGCGTTTCGACATAGAAAGTGCATGCTTCTATGCGCCACCTGGAAACGACAGCA 1380
Qy      1337 TTAAAGAAAGTACCTGCGCTCTTATGATCTATTTGATTCACGCGCTTACCGACGAGG 1396
Db      1381 TCGTAAAGTCCAGCGCTCTTTTAGCGCTTCCCATGTCTATCGGTATGGACTACCTCT 1440
Qy      1397 TCGATATGGGCTTCCGCGTATGAGCTCAAGCTGCTGCTTGAACCA---ACTGTCAACCCG 1453
Db      1441 TGCCCATGTGCTGAGGACAGAAAGTGAATTTGGCATTTCAACCAAGAGAGAGAAAAAC 1500
Qy      1454 CACCGGTATTAACAATGCGGATGTGAGCATCTGCGTGGCTTACAGCAAGAGCTGAG 1513
Db      1501 TGTGTCAAGTCCCGAGAGAAATTAAGTATGAGAGCCAAAGGCTGTTGAGAGATGCTCAG 1560
Qy      1514 AAGTGTGTCAGCGGAGAGATGTCAGAGAGCCCTGCGCACTTGTCTCCCTTAATAGAAA 1573
Db      1561 AGGAATCAGAGCGGAGAAAGCTCGAGAAAGCATCCCACTTGTGTGCAACAAAGGTA 1620
Qy      1574 AAGAGACCTAGAGGAG---AAGTAGACTCATTTATCAAGAGGACAGAGAG 1624
Db      1621 TCGAGGAGCTGCGGAAAGTTGTCTGCGAAGTGAAGGGGCTCCAGCGGACACCGGAGC 1680
Qy      1625 GTACGCTGAGACACACAGAGACATCAGGGTGACAAAGTTACCCAGCGGAGAGAGA 1684
Db      1681 CACTGTGCAAAACCCGCGCGGTATGAAAGTAAATACCTCAAGCAAAATGACGTATGA 1740
Qy      1685 TTGGGTCTTACGCTATCTTCAACCCAGGGGCTTAAATGATGAAAAATTTGGCGTGA 1744
Db      1741 TCGGACAGTATATGTTGTCTGCGCGATCTCTGTGTGAGAAAGCACTAACTCCACAG 1800
Qy      1745 TCCACCATTTGGCGGAAACAAGTACTGTGATGATCTCAAAAGTAGGAGGAGATACA 1804
Db      1801 CACACCGGTAGACAGACAGTTTAAGATCAATACGCACTCCGGAAGATCAGAAAGTATG 1860
Qy      1805 AAGTCAGCCATACACAGGTAAGGTCAITGTACAGAAAGGACGCGGTCCCTGTTCAG 1864
Db      1861 CAGTCGAACATACGAGCTAAAGTATGATGCGAGCAGAAAGTGCATGATGCGCAG 1920
Qy      1865 ACTTCCAGGATTTAGAGAGAGCGCTACATCGTTTTCACAGAGAGGAGTTTCTTAAACA 1924
Db      1921 AATTTTTAGCATGTAGTAGAGCGCCACGCTTGTGTACAGAAAGAGATTTTGTGAACC 1980
Qy      1925 GATACCTGCACCATCGCATCAACGAGAGGCGCTTAAACATGACGAAAGTACTATA 1984
Db      1981 GCAAGCTGTACATATGCGATGACGCTCCCGCTAAGAAATCAGAAAGAGAGACATACA 2040
Qy      1985 AGACTGTAAAGACTCAGAGACACAGCTCAGAAATCGTCTTGATATTTGACGCAAGAA 2044
Db      2041 AGGTTTAAAGGAGAGCTGCGAAGAAACAGAGTACGTGTTGACGTGACAAAGAGCAT 2100
Qy      2045 GTGTAAAGCGAAGAGCGCAGGTCCCTTGTGCTTACCGGTGATCTGTATATCACCAT 2104
Db      2101 GCGTTAAGAGAAAGAGCTCAGAGACTTGTCTTTTGGGAGAACTGACCAACCCGCGCT 2160
Qy      2105 TTACAGAGTTTGGTACAGAGTCTCAAGACAGCAGCAGACCTCACAAGTCCCAA 2164
Db      2161 ATCAGAACTAGCTCTTGAGGAACTGAAAGCTGACCCGCGGTCCGTACAAAGTTGAAA 2220
Qy      2165 CCATCGAGTCTATGAGTGCAGGTTTCAAGTAAATCTGAAATCATCAAAAGCGCTGTGA 2224

```

```

Db      2221 CAATAGAGTATAGGACACACAGATCGGCAAGTACAGTATATCATCAAGTCACTGTCA 2280
Qy      2225 CTAAAGAAAGATCTGTTGTAGTGCAGAGAGAAATCTGCGCAAAATCATCAGGATG 2284
Db      2281 CCGCAGTGAATCTTGTATACAGCGGAGAAAGAAATCTCCCGGAATTTAGAGCCGACG 2340
Qy      2285 TAAAGAGATAGACGATATGATGTTGCTGCTAGACCTGTGATCTAGCTCTTAAATG 2344
Db      2341 TGTACGCGCTGAGGGGATCATCATGCTGCAAGACAGTGAAGATGCGTTATGCTCAAG 2400
Qy      2345 GGGTTAAGCACCCCGTTAACACTGTGATCAATGATGAGCAATTTGCTCCATGACAGGA 2404
Db      2401 GATGTCACAAAGCGGTGAAGTGTATGTCGAAGCGTTCCGCTGCACACAGAG 2460
Qy      2405 CGCTGCTGGCACTGATGGCATGTGTCAAC---TAAAGAGTATTTGGGCGGAGC 2461
Db      2461 CACTACTGCTTGTATGCAATGTGACACCCGTAAGAGATGATCTATGCGGAGACC 2520
Qy      2462 CAAAACATGCGGCTCTTTTAACTGATGTGCTGAAAGTACATTTTAACT----- 2512
Db      2521 CTAAAGCAATGGGATTTCTTCAATGATGCACTTAAAGTATCACTTCAACACCTGAAA 2580
Qy      2513 ATGACATATGCACTGAAGTATACCATTAAGATCTTATAGAGTGCACACACTGTAA 2572
Db      2581 AAGCATATGTACCAAGACATTTACAAAGTTTATCTCCACGTTGCACACAGCCAGTCA 2640
Qy      2573 CCGCATGTCTCCAGCTCTTCTAGAGCAAGGAAATGAAGGATTAACCATGTGCTG 2632
Db      2641 CCGCTATTTATGACACTGCTCATTTACAGTGAAGAAATGAAGAACCAAAACCGTGAAA 2700
Qy      2633 ATAAATCATATAGATACCAAGGAGCACAAAAGCGCAAAAGATGATCTGATTTAA 2692
Db      2701 AGAATCATGAATATGACATTAACAGGGGCAAGAGCGGAAGCCAGGGGACATCATCTGA 2760
Qy      2693 CTTGTTCAAGAGATGGGTGAAGACGTACATGATTTGCTCAAAATACAGAAATCATGA 2752
Db      2761 CATGTTCCCGGGGTGTTAAGCACTGCAAAATGCACTATCCCGGACATGAGGTAATGA 2820
Qy      2753 CTGCGGTGATGCGCAAGGACTTAACGCGGAAAGCGTTTATGCTGACAGTACAAAGTCA 2812
Db      2821 CAGCGCGGCTTCAGAGGCTTAACAGAAAGAGATATGCGCTCGCGCAAAAGTCA 2880
Qy      2813 ACGAATTCACCTTACTCTGACAGCTTCTGAGACGTGAACGTTTACTTACACGACAG 2872
Db      2881 ATGAAACCCGCTGTATGCGATCATCATGAGGATGTGAAGTGTCTCACCCGACCTG 2940
Qy      2873 AAAAAGCATTTGTCTGAAGACGTAGCTGTATCTCTGATTAAGACACTTACACTA 2932
Db      2941 AGGACAGCTATGATGAAAACTTTACAGGGCGCACCATGATTTAACAGCTCACTAACG 3000
Qy      2933 AATATCCCGGGGATTTACAGGCTTCAATTGACGACTGCGAGCGGAAACAGCGCATTA 2992
Db      3001 TACCTAAGAAATTTACAGGCCACATCGAGACTGGAAAGCTGAACCAAGGAAATTA 3060
Qy      2993 TGGCAGCGTCTTGTATAGCCGACAGCTGATGTGTTCCAGAAATAGGTGAACGTCT 3052
Db      3061 TTGCTGGATTAACAAGTCCGCTCCCGTACCAATCCGTTACGTGCAAGACTAACGTTT 3120
Qy      3053 GCTGGCGAAGGCTTTAGAGCACTTTGGCAGCGCCAAACATTTGTGTCAGACAGAC 3112
Db      3121 GCTGGGGAAGGCACTGAAACGATACGTGCGACAGCGCGGTATGTACTTACCGGTTGCC 3180
Qy      3113 AGTGG---GAGAGCTTACCCATTAAGCATGACAGAGGATCTACCTGAAATGCGAC 3169
Db      3181 AGTGAAGCGAGCTGTCCACAGTTGCGAGTGAACCAACACTCGGCCATCTACGCTCT 3240
Qy      3170 TGAATCTTTTTCAGCAGAGTTCTTTGAGTGAAGCTGACAGTGGTATTTTCCGCTC 3229
Db      3241 TAGACGTAATTTTGATTAAGTTTTCGCGATGACCTTGACAAAGGGGCTGTTTCCAAC 3300
Qy      3230 CTACCGTGCACCTTACTTACAGGAT-----CAGCATGGGATA 3268
Db      3301 AGAGCATCCGTTAAGTACCATCTGCGACTGAGGAGCCAGTAGCTCATTTGGGACA 3360

```

QY 3269 ACTCCAGGGAAGACATGTATGGCTTAAATAGAGGTAGCAAGAGTTGTACGGC 3328
 DB 3361 ACAGCCAGGAACAGCAAGTATGGGTACGATCAGCCGTTCCGCCAAGCTCTCCGTA 3420
 QY 3329 GATATCCGTGCATCAAAAAGCGTTGACACAGCGGTAGCTGATATAGAAATATA 3388
 DB 3421 GATTTCCGGTGTTCAGGTAGC---TGGGAAAGGACACAGCTTATTTGACAGGGGA 3477
 QY 3389 CCATCAAGAGACTACTCTCCCAATTAATGTGTTCCATTAATGGCCGGTGGCCCACT 3448
 DB 3478 GATCTAGGATTAATCTCTGACAGCATTAATGTTCCAGTGAACGCAATCTCCCTCACG 3537
 QY 3449 CGTTGATCGTTGACCAAAAGGACAGGGTACAACTGATACAGCGGATCTTATCTAAGA 3508
 DB 3538 CTTTATGCCCGAGACAAAGGAAACAACCGGCCCGGTGCAAAAATTTCTTGACCGAGT 3597
 QY 3509 TGAAGGGCAATCTGTGTGTGTGTGATCGCGGATCTTA-----TCAGATTCGAGG 3558
 DB 3598 TCACACACCACTCCGTACTGTGTGATCTCAGAGAAAAAATTAAGCTCCCAAGAGAA 3657
 QY 3559 AAGAAAGTATGATCATGGGTTCATTCGCCACTAATACATCAGGTGATCTCGATTGG 3618
 DB 3658 TCGAATGATGCCCCGATTTGGCATAGCCGCGCAGATTAAGAACTACAACTGGCTTTCG 3717
 QY 3619 GGAATACCTAGCATGTGCGTAAATATGACATTAATCTTGTGATTTAGACCCCGTAC 3678
 DB 3718 GGTTCCTCCCGCAGGACGG---TACGACTGTGTTCATTAATTTGAACTAAATAC 3773
 QY 3679 AGGAACCATCACTACCAAGTGGGAGATCAAGCTATCCACACAGCATCTGTAACGTGT 3738
 DB 3774 AGAAACCATCACTTCAACAGTGGGAGACACGGCGGACCTTGAACACCTTTGGCGT 3833
 QY 3739 AAGGTGTCCACCACTGTAACATCTGGCGGAACATGTGTGCTATAGGGTATGGCTTGT 3798
 DB 3834 TCGGCTCTGAATGCTTTAAACCCGAGGAGACCTGTGTGTAACTCTCAACGTTACGCG 3893
 QY 3799 GATCGCGAACCAGAAATATCATCTGCGGTGACAGCGCTCATTTAGTTTACCCTGTTC 3858
 DB 3894 GACCCCAATATGAGAGAGCTGATCAACGCTCTTGGCAAAAATTTGTCAAGATGTCTGA 3953
 QY 3859 TGTCACTTAAGAACACTGCGCGAAATATCTAGGTTCTTCTGTGTTCTTTCGGCAGAG 3918
 DB 3954 GCGAGGCGAGAGTGTCTCAAGCAATACAGAAATGATCTGATTTTCGACAACTAGAC 4013
 QY 3919 AAGCCCAACCAACATGACCAAGACAGACTCGGTGATGCTTGAACAATCTATCAA 3978
 DB 4014 AACAGCCGACACGCAATTCACCCCGCATTTGAATGTGTGATTTGCTCGCTGTAC 4073
 QY 3979 GGGTCAACAGGTACGAGGACAGGAGAGCTCCAGGTTACAGAGTATCAGAGGTACATT 4038
 DB 4074 GAGGTTACAGAGAGAGAGTGTGAGCCGACCGTGTATCCGTAATAAGGAGAACTTT 4133
 QY 4039 AGCAAGACCGGTGACCAAGCTATCTGTAATGCTGTAAATAGCAAGGTCAACAGGTTCC 4098
 DB 4134 GCTGATTTCAAGAGAGAGAGTGTATCATATGACCAATCCATGGGACAGACAGAGAA 4193
 QY 4099 GGAAGTGTGGGTGACCTGTACCGAAATGGCCGCTGTTTGTATGACAGCCATAGCT 4158
 DB 4194 GGAAGTGTGGGTGACCTGTATTAAGCTGTGGCGAAAGTTTCCAGATTCAGCCCAAG 4253
 QY 4159 GTCCGAGCGGTAGACT---TGTGAAGCAGAAACGCTCATATCATGCTGTAGAGCC 4215
 DB 4254 ACAGGTACCGCAAACTGACTGTGTGCAAGAAAGAAATGATACAGCGGTGTGGCTT 4313
 QY 4216 AATTTTCTAAGATCCGGAACCGGAGGCGCACTTAAGCTGACGCTCTACATGAGC 4275
 DB 4314 GATTTCCGGAACACCCAGAGGAGAGCCCTGAATTTGCTCAAAAGCCCTACATGCA 4373
 QY 4276 ATAGGTCTATCTGTAAGCTGAGGAGATTAACAAATATCATGTAACGCTACTGTCAAC 4335
 DB 4374 GTGGAGACTTAAGTAATGAACATTAATATCAAGTGTGCGCATCCCACTGATCTACA 4433

QY 4336 GGCATCTATCTGTGGCAAGATGAGTATGATCAATTCATCACCTGTTCATCTGCT 4395
 DB 4434 GGCATTTACGAGAGCCGGAAGAACCGCTTGAGTATCACTTAATCGCTTGACACCGCG 4493
 QY 4396 TTGCACCTAGGATGCCGATGTCAACATATTTCTTGATTAACATTTGGAGACGAG 4455
 DB 4494 CTAGACAAATCTGATGCGGACGTAAACATCTAATCTCCGTGAATMAAGTGAAGAAAGA 4553
 QY 4456 ATTAATCGAGGCATTCACCGCAAGAAAGCGTC---GAAATCTGATGATGACAAGCA 4512
 DB 4554 ATCGACCGCGGTCTCAACTTAAGAGTCTGTATCTGAGCTGAAGGATGATTTGAG 4613
 QY 4513 GTAGCATTTGACTTGTGTCAGGTTCCACCAACAGCTTTTGGCAGGACAGACAGTTAC 4572
 DB 4614 ATCGACAGAGATTATGATGATTCATCGGACATTCGCTGAAGGAAAGAAAGGATTTC 4673
 QY 4573 TCCGTCAATGAGGCAAGTTGATTTATCACTCTGGAAGGTACAGATTCCATGACCGCC 4632
 DB 4674 AGTACTACAAAGAAAGGATGATTCGTATCTTGAAGGACACCAATTCATCAGCAGCA 4733
 QY 4633 AAGCATTTGCGGAATCCATGCAATGTGGCCCAACAAATCTGAGGCTTAATGAGAGATT 4692
 DB 4734 AAGATATGGCGGAGATTAAGTCTCTGTTCCAAATGACAGAGAAAGCAAGCAACTG 4793
 QY 4693 TGTCTGTACATCTTGGGAGAGATGTGTCAAGCATCCGCTCCAAATGCCCAGTAGAGAG 4752
 DB 4794 TGTGCTTACATATTTGGGAGAGACATGGAAGCAATCCCGGAAATTTGCCGTTCAACAC 4853
 QY 4753 TCAAGGCGTGTCTCTCAACCTTCAACACTTCATGCTGTGTATTAATGATGACGCT 4812
 DB 4854 AACCGGTCTAGCCCGCAAAAACGCTGCGTCTCTGTATGATGATGATGATGATGATG 4913
 QY 4813 GAGCGCTATACAGTTTGGCTCTGCGCAAGAAAGAAAGAGTTCGCGGTATGCTCATATTC 4872
 DB 4914 GAAAGGTCTCAAGCTCAGAAAGCAATACGTCAAAGAGTTACAGTATGCTCTCCACC 4973
 QY 4873 CTGTTGCCGAAGTACAGATCAACAGCGCTGCAAGAGTACAGTACAGCAAAACAGTCTG 4932
 DB 4974 CCCCTTCAAGTACAAATCAAGATTTTCAAGAGTTTCAAGAGTTTCAAGTACCAAAATGATCTG 5033
 QY 4933 TTTTCAGGCGTGTACCAACGCGCTGTACACCCCGAAGATACGCGGAA 4980
 DB 5034 TTTAACCGCATATCCCGCATGCTGTCCCGCCGTAAGTACATGAA 5081

RESULT 11
 US-09-102-248-1
 ; Sequence 1, Application US/09102248
 ; Patent No. 6008035
 ; GENERAL INFORMATION:
 ; APPLICANT: Johnston, Robert E.
 ; APPLICANT: Davis, Nancy L.
 ; APPLICANT: Simpson, Dennis A.
 ; TITLE OF INVENTION: System for the In Vivo Delivery and
 ; TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Bell Selizer Park & Gibson, P.A.
 ; STREET: 1211 East Morehead Street
 ; CITY: Charlotte
 ; STATE: No. 6008035th Carolina
 ; COUNTRY: USA
 ; ZIP: 28234
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; FILING DATE: US/09/102,248
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/801,263
FILING DATE: 19-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-147
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11663 base pairs
type: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 60..7559
FEATURE:
NAME/KEY: CDS
LOCATION: 7608..11342
US-09-102-248-1

Query Match 26.7%; Score 1333.2; DB 3; Length 11663;
Best Local Similarity 56.1%; Pred. No. 0;
Matches 2820; Conservative 0; Mismatches 2128; Indels 80; Gaps 13;

23 ATATGAAAGATTCACGTTGACTTGAAGTGTGACAGCCGCTATGTCAAGTCGTTACAC 82
Db AGAAGCCAGTAGTTAGCTAGACCTAGACCTCCTGAGTCCGTTTGTCTGCACTGCAAA 123
Qy 83 GGAAGTTTCCACATTTGAGATGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 142
Db 124 AGAGCTTCCGCAATTTGAGTAGTAGAGCAGAGGTACTCCAAATGACCATCTAATG 183
Qy 143 CCAAGAGCTTTCCGATGTGCAACAAAGCTCATTTGAGAGGAGGAGGAGGAGGAGGAG 202
Db 184 CCAAGAGCTTTTCCGATGTGCAACAAAGCTCATTTGAGAGGAGGAGGAGGAGGAGGAG 243
Qy 203 TTATCTTGAACATTTGAGAGTGGCGCCCTGACAGATGCAATTCATCAACCGCTATCAT 262
Db 244 CGATTTTGAACATTTGAGAGTGGCGCGCCCTGACAGATGCAATTCATCAACCGCTATCAT 303
Qy 263 GTATCTCCCTATGATTAAGGCTGAAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 322
Db 304 GCGTTTCCCTATGATTAAGGCTGAAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 363
Qy 323 TTAAAGAAA-----GTGACATTTACCGACAAAGCATAGCTCTTAAGGCGGAGAGCTGC 376
Db 364 TGGCGGAAAAGCATGTGAAGATTACAAACAAAGAACTTGATGAGAAAGTCAAGAGACTCC 423
Qy 377 TGAAGTCACTCAACACAGAGCGAAGACTCATCTCTGTGTATGACACAGAGCGCA 436
Db 424 GGAACCGTACTTGATACACCGGATCTGAAGCGCATCACTCTGCTTCCAAACATGTTA 483
Qy 437 CGGTAGAGTCTTGAAGTGTACAGATACCAAGATGTGTAGGAGCTCCATCAACCGA 496
Db 484 CTTGCAACAGCGGTGCGAGTACTCCGTATGAGAGAGTGTACA--TCAACGCTCCCG 540
Qy 497 CATCAATCTACCAACAGCGCTTAAGAGGATTAGAACAAATTACTGTAGTAGCTTTGACA 556
Db 541 GAATCTATTACCAACAGCGCTTAAGAGGCGTGGCGAGCCCTGTACTGATTTGGCTTTGACA 600
Qy 557 CGAGCCCTTTTATGTACAAAACATGCGAGTTCTTACCTTACTTACCAACGAACTGGG 616
Db 601 CCAACCCAGTTCATTTCTCGGCTATGCGAGGTTGCTGATCCCTGATACCAACCAACTGGG 660
Qy 617 CTGACGAGAGATTTGAAGACAGTAACTTGGCTCTCGTAATCTCAATTTTGAAGAGA 676
Db 661 CCGAGCAAAAAGTCTTGAAGCGGTAACTCGAGACTCTGAGACAAAGCTGAGTAGG 720
Qy 677 GCAAGCTTGGAACCTCAATCTTAGAGAGAGAGGCTCCAACTACTAATAGATCA 736

Db 721 GCAAGACAGAAAAGTTGCTCATATAGAGAGAGAGGTTGAAGCCCGGGTCAAGGGTTT 780
Qy 737 TATTTCTGGTTTGTTCACAAATCTACACAGAAATAGATACATGTTACATAGCTGGATC 796
Db 781 ATTTCTTCTGTTGATTCACCTTTTACCCAGAACACAGAGCAGCTTGCAGAGCTGGATC 840
Qy 797 TTCCAAAGTGTTCACCTTAAAGAAAGCTTAATCTTACAGAGTGAATGTGGACCAATTG 856
Db 841 TTCCATGGGTGTTTCACTTTAAGAGAAAGAGCTGTACATTTGGCCGTGTGATACATGG 900
Qy 857 TCAAGTGAAGGATCGTATCAAAAAGATTAACGATACAGCCAGAGCTATACGTTAAG 916
Db 901 TGAGCTCGAAGGCTAGCTAGTGAAGAAATACACATCACTCCCGGATACAGGAGAAA 960
Qy 917 TTGAGAACTTGGGCTCCAAATGATCGCAGAGGTTTCTTGAATGGAAGTCAAGATA 976
Db 961 CCGTGGGATACGGGTTTCAAAACAAATAGCAGAGGCTTTTGTCTATGCAAAAGTTACGATA 1020
Qy 977 CGCTGCGGCGAGAGGGTTTCTTTTGTGTGTAGCTATGATACAGCCACACTTTGGC 1036
Db 1021 CAGTAAAGAGAACGGGTATCGTTCCCGGTGTGACGTATATCCCGCCACCATATATGG 1080
Qy 1037 ATCAGATGACAGGATTTCTGGCAACTGACGTTAGTGTATGACGACCAAAAATCTATTGG 1096
Db 1081 ATGAGATGACCGGATATATGCGACGAGATATCTCACTGACGATGACAAAATCTTCTGG 1140
Qy 1097 TTGGGCTTCAACCAAGATTTGTCTCATATGTAGAGAGCAAAAGAAATATCAACATATGC 1156
Db 1141 TTGGGCTTCAACCAAGATTTGTCTCATATGTAGAGAGCAAAAGAAATATCAACATATGC 1200
Qy 1157 AGAATCTATATTATACAGATGATGCGCCAGAGGCTTTTCCAGTGGGCGCGTGAACATGCTG 1216
Db 1201 AAAATTTACTTTCTGCTCAATATGTCACAAAGGTTTACCAATATGGCCCAAGAGCGCAAG 1260
Qy 1217 CCGACTTGAACAGAGAGAAAGAACTAGAGGCTGCGGAGGCGCACTTTATCTATGGCTGCT 1276
Db 1261 AAGATCTTGAACATTAAGAAATATGCTGGCACAGAGAGCGCAAGCTTACATATGCTGCT 1320
Qy 1277 GCTGGGCTTTCAGAGCCCAAGAAATCATATCTTCAAGAAAGCTGTATACGAAACAA 1336
Db 1321 TGTGGGGTTTTCGACATTAAGAAAGTCACTGTTATATGCGCCCACTGGAACGAGACCA 1380
Qy 1337 TTAAAGAAAGTACCTGCGGCTTTTGAATCTTGAATTTTCAAGGCTTATACAGCCAGCGGC 1396
Db 1381 TCGTAAAGTCCAGAGCTCTTTTATAGCGCTTTCCCAATGTATCCGTATGAGACTATCTT 1440
Qy 1397 TCGATATGAGGCTTCCGCTGATGAGCTCAAGCTGCTGTGAACCA--ACTGTCAACCCG 1453
Db 1441 TGCATATGCTGCTGAGGCAAGATGAAATTTGACATTAACCAAGAAAGAGGAGAAAAC 1500
Qy 1454 CACCGGCTATTAACAATGCGGATGTGAGCATCTGCGGTGCTTACAGCAAGAAAGCTGAG 1513
Db 1501 TGTGCAAGTCCCGAGAGAAATGATTAATGAGAGGCAAGGCTGCTTTCAGAGATGCTCAGG 1560
Qy 1514 AAGTGTGAGGAGGAGAGATCAGAGAGCCCTGCAACCTTGTCTCTGAAATAGAAA 1573
Db 1561 AGGATTCAGAGCGGAGAAAGCTCCGAGAGCACTTCCACATTAATGTGAGCAAGAAAGTGA 1620
Qy 1574 AAGAGCCGTAGAGGAG-----AAGTAGCTCATTAATGCAAGAGGAGGAGGAGGAG 1624
Db 1621 TCGAGGAGAGTGGCGAAGTTGTCTGCAAGTGTGAGGAGGAGGCTCCAGGCGGACACCGGAGCAG 1680
Qy 1625 GTAGCGTGAAGACACACAGAGACATCTAGGCTGACAAAGTTTACCCAGGAGGAGAGAGA 1684
Db 1681 CACTCGTGAAGAACCCCGCGGTCTATGAGATTAATCACTCAAGCAAAATGACCGTATGA 1740
Qy 1685 TTGGGCTTATAGCTATCTTCAACCCAGGCGGATTAATGAGAAACTGGCGTGA 1744
Db 1741 TCGGACAGTATATCTGTTGTTCTCCGCAATCTCTGTGCTGAAGAGCTTAACTGCAACAG 1800
Qy 1745 TCCACCCATTTGGGAGCAAGATCTGTATATGACTCAAAAGTATGGGACAGGAGATACA 1804

Db 1801 CACACCCGCTAGACAGACGAGTTAAGATCAATTAACGCACTCCGAAAGATCAGAAAGGTATG 1860
QY 1805 AAGTCAGAGCATATACCAGCGGTAAAGTTCATTGTATCAAGAAAGGAGCGGCGGTCCGTGTTCAAG 1864
Db 1861 CAGTGGAAACATATACAGCGCTTAAAGTACTATGCTCCAGCAGAAAGTGCCTGATACATGGCCAG 1920
QY 1865 ACTTCAGAGCATTTAGTGAAGACCGGTACGATCGTTTCAACGAGAGGAGTTCGTAAACA 1924
Db 1921 AATTCTTAGCACTGAGTGAAGACCGCAAGCTTGTGACAAAGAAAGAGTTCGTGAAC 1980
QY 1925 GATACCTGACCACTATCCGATCAACGAGAGCCGTAAACACTGACGAAGTACTATA 1984
Db 1981 GCACCTGATCCATATTTGCTACGACGCTCCGCTAAGATACAGAAAGAGGAGCTATCA 2040
QY 1985 AGACTGTAAGACTCAGACACAGACTGAGATAGCTTCGATATATGACGACGAAAGT 2044
Db 2041 AGGTTACAAAGGCAAGCTCCGCAAAACAGAGTACGTGTTTACGCTGGCAAGAAAGCAT 2100
QY 2045 GTGTTAAGCGAAGAACGACGAGTCCCTGTGTCTTAACCGGTATCTGTAGATCCACAT 2104
Db 2101 GCGTTAAGAGGAAGAGCTCAGAGACTTGTCTTTCCGGAAGATGACCAACCCGCCCT 2160
QY 2105 TTCAAGAGTTTCCGTACAGAGTCTCAGACACAGACAGACACTTCAAAAGTCCAA 2164
Db 2161 ATCAGAACTAGCTTTGAGGACTGAAAGACTGACCCGCGGTCCGTACAAAGTTGAAA 2220
QY 2165 CCATCGAGTCTATGAGAGTCCAGGTTCCAGTAAATCTGGAATCAATCAAAACGCGTGA 2224
Db 2221 CAATAGAGATATAGACACACAGAGTCCGCAAGTCACTATCAATCAAGTCAACTGTA 2280
QY 2225 CTAAAGAAAGATCTGTTGTGAGTGCAGAAAGAAAACTCCGCAAGATCATCAGAGATG 2284
Db 2281 CGGCAAGTATTTGTTACACAGGAAAGAAAGAAACCTCCGCAAGATGAGCGCAG 2340
QY 2285 TAAAGAGATGACAGTATGATGTTGCTGTAGACTGTCCATTCAGTCTTCAATG 2344
Db 2341 TGCTACGCTGAGGGGCACTGACATCAGTCCGAAAGAGTGAATTCGTTATGCTCAAG 2400
QY 2345 GGGTTAAGACCCCGTTAAACACTGTATCATTTGAGAGGCAATTCGTCGCAATGAGGAA 2404
Db 2401 GATGCCAAGAACCCGTAGAGTGTGTATGTTGAGCAAGCGTTCCGATGCCAGAGGAG 2460
QY 2405 CGCTGCTGCACTGATGTCATGCTCAAAAC---TAAAGAAAGTGTATTTGTCGGGAG 2461
Db 2461 CACTACTTGCCTGATTTGCAATGCTCAGACCCCGTAAAGAGTATGATCTATCGGAGAC 2520
QY 2462 CAAACCAATGCGGCTTTCTTAAATGATGTCCTGAAGTACATTTAAAC----- 2512
Db 2521 CTAAACCAATGCGGATTTCTTCAATGATGCACTAAAGGTACATTTCAACACCTGA 2580
QY 2523 ATGACATATGCACTGAAGTACATTAAGATCTTATAGAGTGCACACAGACTGTAA 2572
Db 2581 AAGACATATGTAACAAGACATTTCTACAGTTTATCTCCGAGCTGCAACAGCGAGTCA 2640
QY 2573 CCGCATGCTGCTCAAGCTCTTCTACGACAAAGCATGAAAGAGCGGTATGCCATGTGCTG 2632
Db 2641 CGGCTATTTGATGACACTGATTTACGATGAAAAATGAAAAACCAAAACCCGTGCAAGA 2700
QY 2633 ATAAATCATCATATAGATACACAGGAGACCAAAAGCCGCAAAAGATGATCTGATTTCA 2692
Db 2701 AGAATCATGAAATCGACATTAACAGGGGCGCAAGACCCGAAGCAAGGGGACATCATCTGA 2760
QY 2693 CCTGTTTGAAGAGTGGGTGAACAGCTACAGATTAAGCTTGAACAAATATCGAAATCATGA 2752
Db 2761 CATGTTTCGCGGGGTGTTAAGCACTGCAATGCAATCCGAGCATGAGGTAAGA 2820
QY 2753 CTGCGGCTGCACTCGAAGGACTTACGCGGAAAGCGTTATGCTGTCAAGTCAAAAGTCA 2812
Db 2821 CAGCGCGGCGCTCAACAAGGCTTAACCAAGAAAGAGTATATCCGTCCGCGCAAAAGTCA 2880
QY 2813 ACAGAAATCCACTTACTGACAGACTTGTGACACAGTGAACGTGTTACTTACGACAG 2872
Db 2881 ATGAAACCCGCTGATACGAGTACATCAGACATGTGAACGTGTTGCTCAACCCGACATG 2940

QY 2873 AAAACGCAATTTGTGGAAGACGCTAGCTGTGATCCCTGATTAAGACACTTAACAGCTTA 2932
Db 2941 AGGACAGGCTATATGAAAACTTTACAGGGGAGACCATGATTTAAAGCTCATTAAG 3000
QY 2933 AATATCCCGGGATTTACAGGCTTATGACAGTGGCGAGCGCAACGACGCCATTA 2992
Db 3001 TACTTAAAGAAATTTTACAGGCACTATGAGAGCTGGGAGCTGAAACCAAGGAAATTA 3060
QY 2993 TGGCAGCGGTTCTTATTAAGCCGACAGACTGATGTGTTCCAGATTAAGTGAACGTCT 3052
Db 3061 TTGCTGCATTAACAGTCCCGCTCCCGTACCAATCCGTTACGTGCAAGACTTAACGT 3120
QY 3053 GCTGGGCGAAGCTTTAGAGCCAGTCTTGGCCAGCGGCAAACTTGTGCTGACGAGACAG 3112
Db 3121 GCTGGGCGAAGCACTGGAACCGATATGCGCCAGCGGTATCGTACTTACCGGTTGCC 3180
QY 3113 AGTGG---GAGAGCTTGACCCCATTCAGCATGACAGAGCGTACTACCTGAAATGGCAC 3169
Db 3181 AGTGAAGGAGGCTGTCCCAAGCTTTGGGATGACAAACCAACTCGGCATCTACGCCCT 3240
QY 3170 TGAATCTTTTGTGACACAGGTTCTTGGATGACCTGGAACGTGGTTATTTCCGCTC 3229
Db 3241 TAGAGTAAATTTGCAATTAAGTTCGTCATGACTGACCTTGAACGCGGCTGTTTCCAAAC 3300
QY 3230 CTACCGTGCACCTTACTTAACAGGAT-----CAGCACTGGGATA 3268
Db 3301 AGAGCATCCCTTAAAGTACATCTCTGCCACTCAGCAGGCGCACTAGCTTATGGGATA 3360
QY 3269 ACTCCGAGGAAAGAACTGATGGGCTTAAATGAGAGTACAAAGAGTGTGACGCG 3328
Db 3361 ACAGCCGAGGAAACAGCAAGTATGGTATGATCAAGCGGTTCCCGCAACTCTCCGTA 3420
QY 3329 GATATCCGTGATCAAAAGCGGTTGACACAGGAGGATGATATTAAGAAATATA 3388
Db 3421 GATTTCCGCTGTTCCAGCTAGC---TGGGAAAGGACACAGCTTATTTGACAGCGGCA 3477
QY 3389 CCATCAAGAGTACTCTCCCAATTAATGAGTTCATTAATGCGCGGTGGCCCACT 3448
Db 3478 GAACTAGATTAATCTCTGACAGCATATCTGTTCCAGTGAACCGCAATCTCCTCAG 3537
QY 3449 CGTTGATGTTGACCAACAAAGGACAGGATACAACTGATCAACAGCGGATTCATTAAGA 3508
Db 3538 CTTATGCTCCCGAGACAAAGGAAACAAACCGCGCCGCTGCAAAATCTTGAACCACT 3597
QY 3509 TGAAGGCAATCTGTGTTGATGCGGATCTTA-----TACGACTTCCAGG 3558
Db 3598 TCAACACACACTCCCTACTTGTGATCTCAGAGAAAAAATTGAAAGCTCCCAACAGAGA 3657
QY 3559 AAGAAAGTAAAGTCCATGGGTCCATGTCCTAATATACATCAGGTGATCTCATTTG 3618
Db 3658 TCGAATGATGCGCCCATGTTGGCATAGCCGGGCGAGATTAAGAACTAACCTGGCTTGC 3717
QY 3619 GGAATACCTAGCACTGTGTAATATGACATTAATCTTGTCAATGTTAGACCCCGTAC 3678
Db 3718 GGTTCCTCCCGAGGACAGG---TAGACCTGTGTTTCAATTAATTTGAACTAAATTA 3773
QY 3679 AAGAACTCATCTACCAACAGTGGAGATACGCTATTCACACAGCATGCTAACGTT 3738
Db 3774 AGAAACCATCACTTTCAACAGTGCAGAACACACGCGGCACTTGAACCTTTGCGGT 3833
QY 3739 AAGGCTGTCACACCTGAAACATGCGGGAACATGTGGGCTATAGGGATGCGGCTGT 3798
Db 3834 TCGGCTCTGAATGCTTTAACCCCGAGGACACTCGTGGTGAAGTCTTACGTTTACGCG 3893
QY 3799 GATCGCGCAACGAAATATCATCTGCGGTGACGCTCATTTTAAAGTTTACCCGTTGC 3858
Db 3894 GACCGCATATGAGGAGAGTATGATCAGCTCTTGCACAAATTTTGTAGAGTGTCTGA 3953
QY 3859 TGTACGCTTAAGAACACTGCGGAAATATCTAGGTTCTTGTGTTCTTGGCAGAGAC 3918
Db 3954 GCGAGGCGCAGAGTGGTCTCAAGCAATACAAATATGATCTGATTTTCCGAACACTAGAC 4013

QY 377 TGAAGTATGTCAACACGAGAGAGACTCATCTCTGTGTATGACACAGAGCGCA 436
Db 424 GGACCGTACTTGATACACCGGATGCTGAAGCGCATCACTGTCTTCCACACAGATGTTA 483
QY 437 CGTGTAGTACTTTGGAGTGTGACATATACCAAGATGTGTAGCCAGTCCATGACCGA 496
Db 484 CCTGMAACCGCGTCCAGTACTCCGTATGCAAGAGAGGTGACA---TCAACGCTCCCG 540
QY 497 CATCAATCTACACAGGCGCTTAAGAGATTAGGACAAATTACATGATAGGCTTTGACA 556
Db 541 GAATATTATACACAGGCTATGAAGGCGTGGACCTGTATGTGATGTGCTTTCACA 600
QY 557 CGACCCCTTTTATACAAAAACATGAGAGGTTCTTACCTACTTACAAACAGATCGG 616
Db 601 CCACCAATCTATGTTCTCGGCTATGCGAGGTTCTTACCTGATACAAACACATCTGG 660
QY 617 CTGACGAGAGATTTGGAAGCAGTAATTTGGCTTGGTAACTGAGATCTTCAGAGA 676
Db 661 CCGACGAAAAAGTCTTGAGAGCGTAAATCGAGCTGACGACAAAGCTGAGTGAAG 720
QY 677 GCAGCTTGGAAAACTCTCAATCTTAGAAGAAAGGCTCCAACTACTAATAAGATCA 736
Db 721 GCAGACAGGAAAGTGTGATATGAGAGAAAGAGTTGAAGCCGCGGTCAAGGCTTT 780
QY 737 TATTCTCGTGTGTTCAACAAATCTACAGAAAGATGATCACTGTTAGGTAGCTGGCATC 796
Db 781 ATTTCTCGTGTGATGACACTTTACCGAAGAACAGGCCAGCTTGACAGACTGGCATC 840
QY 797 TTCCAAAGCTGTTCACCTTGAAGGAAAGTCTTACAGATGATGTGGACCATATG 856
Db 841 TTCCATCGGTGTTCACCTTGAAGGAAAGCACTGACACTGCGGTGTGATACAGTGG 900
QY 857 TCAGCTGTGAAGGGTATGCTCAATAAAGATTAAGATCAAGCCAGAGACTTATAGGTAAG 916
Db 901 TGAGCTGTGAAGGGTATGCTCAATAAAGATTAAGATCAAGCCAGAGACTTATAGGTAAG 960
QY 917 TTGAGAACTTGGGCGTCCACAAATGCAATCGAGAGGTTTCTTGAAGTGAAGTCAAGATA 976
Db 961 CCGTGGGATATGCGCGTTACAAACATAGCGAGGCTTCTTGTATGCAAAAGTACCGATA 1020
QY 977 CGCTGCGGCGAGAGGGTTTCTTGTGTGTATGATGATGATACAGCCACACTTTGCG 1036
Db 1021 CAGTAAAGGAGAAAGGGTATCGTTCCCGTGTGACAGTATATCCCGGACCAATATGCG 1080
QY 1037 ATCAGATACAGGATTTCTGGCACTGACGTTAGTGTGATGATGATGATGATGATGATG 1096
Db 1081 ATCAGATACAGGATTTCTGGCACTGACGATATCTCACTGACGATGACCAAAAACTTCTGG 1140
QY 1097 TTGGGCTCAACCAAGATTTGCTGCAATGGTAGAGCCCAAGAAATCTAACCAATGC 1156
Db 1141 TTGGGCTCAACCAAGATTTGCTGCAATGGTAGAGCTTACGAGAACACCAATACATGAC 1200
QY 1157 AGAACTATCTATTAACAGTGTGCGCCAGGGGTTTTCAGGTGGCGCGGTGAACATCTGTG 1216
Db 1201 AAAATTACTTCTGCAATCAATGCAAGGTTTACGAAATGGGCCCAAGAGCCCAAG 1260
QY 1217 CCGACTTGAAGCAGAGAAAGAACTAGAGGATGCGGAGCGCACTTTACTATGAGGCTGCT 1276
Db 1261 AAGATCTTGAAGAAAGAAAGTCTGGGACACAGAGAGCGCAAGCTTCAATATGAGCTGCT 1320
QY 1277 GCTGGGCTTGAAGCCAGAAATACATGCAATCTAACAAAGGCTGTGAGGAAACAA 1336
Db 1321 TGTGGGCTTTCGCACTAAGAAAGTCACTGTTTATGCGCCACCTGAAAGCGAGACA 1380
QY 1337 TTAAGAAAGTACTGCGCTCTTGTGATGATGATGATGATGATGATGATGATGATGATG 1396
Db 1381 TGTGAAAGTCCAGGCTCTTTTACGCTTTCCTCATGTCATGCTATGAGTACTCTT 1440
QY 1397 TCGATATGAGGCTTCCGCGTATGCTCAAGCTGCTGTAACA---ACTGTCAAAACCG 1453
Db 1441 TGCCATGTCTGCTGAGGAGAAAGATGAATGCGCATTAACAACAAAGAGGAGGAAAAAC 1500

QY 1454 CACCGCTATTATCAATGCGCCAGTGTGAGCATCTGCGTGGCTTACAGCAAGAGCTGAAG 1513
Db 1501 TGCTGGCAAGTCCCGAGAGAAATTAGTTATGAGGGCCAAAGGCTGCTTTCAGAGATGCTCAG 1560
QY 1514 AAGTGTGCGACCGGAGAGATCAGAGAAAGCCCTGCGCACCTTGTCTCCCTGAATAGAAA 1573
Db 1561 AGGAATCCAGAGCGGAGAAAGTCCGAGAAAGCATCCCACTTATGTGACAGCAAAAGTGA 1620
QY 1574 AAGAGCCGTGAGGAGG-----AAGTAGACCTCATTTATGCAAGAGGAGGAGGAG 1624
Db 1621 TCGAGGAGCTGCGAATTGTCTGCGAAGTGAAGGGGCTCCAGCGGACACCGGAGCAG 1680
QY 1625 GTAGCGTGAAGACCAACAGAGAACATCAGGATGACAAAGTTACCAGCGAGAGAGAA 1684
Db 1681 CACTGTGGAATCCCGGGGCTCATGTAAAGATTAATCTCCAGCAATATGACCGTATGA 1740
QY 1685 TTGGGTCTTACGCTATCTTTCACCCAGCGGATTTGAATGTAAGAAAACTGCGCTGTA 1744
Db 1741 TCGGACAGTATATGCTGTCTGCGCATCTGTGTGTAAGAAACGCTAAACTCGCACAG 1800
QY 1745 TCCACCCATTGGCGGAAACAAATCTGTAATGATCAACAAAGTATGAGGAGAGATACA 1804
Db 1801 CACACCGCTAGCAGACCGATTAAATCATATGATCACTCCGAAATCAGGAAAGTATG 1860
QY 1805 AAGTGAAGCCATPACACGCTAAGCTATTGTACAGAGAGAGCGCGCTCCCTGTTCAAG 1864
Db 1861 CAGTGAACCATATGACGCGCTTAAAGTATGATGCGACAGAGAGAGTCCATCATGCGCAG 1920
QY 1865 ACTTCAGACATTGATGAGAGCGTACGATCGTTTTCACAGAGAGGAGTTCTGAACA 1924
Db 1921 AATCTTGAACATGATGAGAGCGCACGCTTGTGTATCAAGAAAGAGATTGTGAACC 1980
QY 1925 GATACCTGACCAACATGCGCAATCAACGAGAGAGCGCTTAAACATGACGAAGATCTATA 1984
Db 1981 GCAAGCTGTATCATATGCTTGCATGACGCTCCCGCTTAAAGATPACAAAGAGAGCTGACA 2040
QY 1985 AGACTGTAAAGACTCAGAGACACAGTCAAGATAGCTTGTGATATGACCAAGCAAGT 2044
Db 2041 AGCTTACAAAGCAGAGCTCCAGAAACAGATGATGCTTGTGACGTGACAAAGAGCAT 2100
QY 2045 GTGTTAAGCGAAGACGAGAGTCCCTTGTCTTAAACCGGTGATCTGTAGATTCACCAT 2104
Db 2101 GCGTTAAGAAAGAAAGAGCTTCAGAGACTTGTCTTTCGGGAGAACTGACCAACCGGCT 2160
QY 2105 TTCAAGAGTTTGGTGAAGAGTCTCAAGACACGACGACACTCAACAAAGTCCAA 2164
Db 2161 ATCAAGAACTAGCTCTTGAAGGACTGAAGACTCGACCGCGGTCCGTTACAAAGTTGAAA 2220
QY 2165 CCATCGAGTCTATGAGATGCGCAGGTTCAAGTAAATCTGAATCATCAAAAGCGCTGTA 2224
Db 2221 CAATGAGATATGAGCAGACACAGAGATGCGCAAGTCACTATCATCAATCACTGTCA 2280
QY 2225 CTAAGAAAGATCTGTTGTGAGTGCAGAGAGAAATCTGCGCAGAAATCATCAGGGATG 2284
Db 2281 CCGAGCGATCTTGTATACAGCGGAAAGAAAGAAATCTGCGGAAATTTGAGCGCAG 2340
QY 2285 TAAAGAGATGAGACATGATGATGCTGCTGATAGACTGTGATGATGATGATGATGATG 2344
Db 2341 TGCTACGCGCTGAGGGGCTGACAGATCACTGTGAAGACAGTGAATTCGTTATGCTCAAG 2400
QY 2345 GGGTTAAGCACCGGTTAACAATCTGTATCAATGATGAGAGCAATTTGCCATGACAGGA 2404
Db 2401 GATGCAACAAAGCGGTGAAGATGCTGTATGTGACAAAGCTTCCGGTCCACGAGAG 2460
QY 2405 CGCTGTGCACTGATGCTATGCTCAAAAC---TAAAGAAAGTGTATTTGTCGGGAGC 2461
Db 2461 CACTACTGCTTGTATGCAATGCTGACAGACCCCGTAAGAAAGTGTATGATGCGGAGAC 2520
QY 2462 CAATAACATGCGGCTTTTAAACATGATGCTCTGAAGATCAATTTTAAAC----- 2512
Db 2521 CTAAGCAATGCGGATTTCTTAAACATGATGCAATTAAGATCAATTTTCAACACACCTGAA 2580
QY 2513 ATGACATATGACATGAAGTATACATAAAGCATCTCTAGAGAGTGCACAGACTGTAA 2572

Db 2581 AAGGCAATGATCCAAAGCAATTTCTACAAAGTTTATCTCCGACGTTGACACAGCAATCA 2640
Qy 2573 CCGGCATGCTCTCCAGCTCTTCTTACGACAAAGCAATGAGACGGTTAACCATGCTG 2632
Db 2641 CGGCTATTGTAACACATGATACGATGGAATAAATGAAACCAACAAACCCGTCAGAA 2700
Qy 2633 ATAAATCATCATGATATACCAAGGACCCCAAGCCGACCAAGATGATCTGATTTCTAA 2692
Db 2701 AGAACATGAAATGACATTAACAGGGGCCAAGAGCCGAAGCCAGGGGACATCACTCGA 2760
Qy 2693 CCTGTTCAAGAGATGGGTGAAGACGTAAGATTGACTACAAATAATCAGCAATCAATGA 2752
Db 2761 CATGTTTCCGGGGTGGGTTAAGCAATGCAATGCAATCCGGACATAGAGTTATGA 2820
Qy 2753 CTGCGCTGATCGCAAGACTTAACGGGAAAGCGTTTATGCTGTCAAGTCA 2812
Db 2821 CAGCCGGGCTTCAAGAGGCTAACAGAAAGATATGCGTCCGGCAAAAGTCA 2880
Qy 2813 ACGGAATCCACTCTACTCGGACACTTGTGACAGTGAAGTGTACTTACAGCAAG 2872
Db 2881 ATGAAACCCGCTGACGCAATCAATCAGACATGTAACGTGTGCTCACCCGCACTG 2940
Qy 2873 AAAAAGCATGTGTGGAAGACGCTAGCTGTGATCCCTGGATAAAGCACTTACAGCTA 2932
Db 2941 AGGACAGCTGATGATGAAACTTTACAGGGGACCATGATTAAGACGCTCATACG 3000
Qy 2933 AATAATCCCGGGATTTACAGGCTTCAATGAGACATGCGAGCGGAAACAGACCATTA 2992
Db 3001 TACCTAAAGGAAATTTTCAGGACCAATCGAGGACTGGGAAGCTGAAACAAAGGAATTA 3060
Qy 2993 TGGCAGCGCTTCTGATTAAGCCGACAGAGCTGATGTGTTCAGAAATGAAGTGAACGCT 3052
Db 3061 TTGCTGCAATTAACAGTCCCGCTCCCGTACCAATCCGTTACGCTGCAAGCTAACGTTT 3120
Qy 3053 GCTGGGCGAAGCTTTAGAGCCAGTCTTGGCCACGGCCAACTGTGTGACGAGACAGC 3112
Db 3121 GCTGGGCGAAGCACTGAAACGATACAGGCGGCGGATGTACTTACCCGTTGCC 3180
Qy 3113 AGTGG--GAGAGCTTGCAACCATTTCAAGATGACAGAGCTACTCACTGAATGCGAC 3169
Db 3181 AGTGAAGGACCTGTTCCACAGTTTGCGGATGACAAACCACTCCGCTCATACGCT 3240
Qy 3170 TGAATCTTTTGAACGAGTCTTTGAGTAGACCTGAGAGGGGTATTTTCCGCTC 3229
Db 3241 TAGACGTAATTTGATTAAGTTTTCGCAATGACTTAAACAAGGCGCTGTTTCCAAC 3300
Qy 3230 CTACGTCGCACTTACTTACAGGGAT-----CAGCACTGGGATA 3268
Db 3301 AGAGATCCGTTAAGTACCATCTGCGACTCAGGAGGCACTACTCATTTGGAGCA 3360
Qy 3269 ACTGCCAGGAGAAACATGATGGGCTTTAATAGAGATGACAAAGATTTGTCAGGC 3328
Db 3361 ACAGCCCAAGAAACAGCAAGTATGGGTACGATCACGCCGTTGCCCAACCTCTCCCTA 3420
Qy 3329 GATATCCGTCATCAAAAGCGGTTGACACAGGAGGGTATGATTAAGGATATA 3388
Db 3421 GATTTCCGGTTCAGCTAGC---TGGGAAAGGACACAGCTTATTTGCAAGGCGCA 3477
Qy 3389 CCATCAAGAGACTACTCTCAACATTAATGTGTTCATTAATGCGCGTTGCCCACT 3448
Db 3478 GAACATAGATTAATCTCTGACACGATTAATTGTTCCAGTGAACGCAATCTCCCTCAG 3537
Qy 3449 CGTTGATGTTGACCAACAAGACAGGGTACATGATCAACGCGGATTTCTATTTAAGA 3508
Db 3538 CTTTATGTCCTCCGAGACAAAGGAAACAAACCGGCGCTCGAAATAATTTCTTGAACCGT 3597
Qy 3509 TGAAGGGAATCTGTGTTGATCGGATCTTA-----TGAGATTCAGAGG 3558
Db 3598 TCAACACCACTCCGTAATGATCTCAGAGAAAAAATTAAGCTCCCAACAAGAA 3657
Qy 3559 AAGAAATGATGATCGATGGGCTCATTTGCCCACTAATACCATCAGGTGATCTCGATTTG 3618

Db 3658 TCGAATGATCGCCGATGTCATAGCCGGCCGACATGAAGAACTCAACCTGCTTTCG 3717
Qy 3619 GGAATACCTTATGATGCTGCTTAATATGACATTAATCTTGTCAATGTTAGACCCGTAC 3678
Db 3718 GGTTCGCGCGAGGACG---TACAGCTGTGTATCAATATTTGAACTTAATATAC 3773
Qy 3679 AGGAACCATCACTACCAACAGTGCAGAGATCAAGCTATCCACCAACAGCATGCTAAGTGT 3738
Db 3774 AGAAACCATCACTTTCAACAGTGCAGAGACCAAGCGGCGACCTTGAAAACCTTTGCGGT 3833
Qy 3739 AAGGCTGTCCACCACTGAACTGCGGGAACATGTGTGCTTATAGGATATGGGCTTGTCT 3798
Db 3834 TCGGCTCTGAATGCTTAAACCCCGAGGACCTCGTGTGTAAGTCTACGCTTACGCC 3893
Qy 3799 GATCGCCCAACGAGAAATATCATCACTCGGTGGACCGCTCATTTAGGTTTACCCGCTC 3858
Db 3894 GACCGCAATGTAGAGCGATGACCGCTTGTGCAAGAAATTTGTCAAGTGTCTGCA 3953
Qy 3859 TGTGAGCTTAAGAACTATGCGGAAATATCTGAGGTTCTTCTGTGTTCTTCCGCAAGAC 3918
Db 3954 GCGAGCCAGAGTGCCTCTCAACCAATACAGAAATGATCTGATTTCCGACAACTAGAC 4013
Qy 3919 AACGCAACCAACACATGACAGAGACAGCTCGGTGTACTGTCACAAATCTATCAA 3978
Db 4014 AACGCGCACACAGCAATTCACCCGCACTATTTGAATGTGTGATTTGCTCGGTAC 4073
Qy 3979 GGTTCACACAGTACAGAGGAGAGCTCAAGCGTACAGAGTATCAGAGTGAATCATT 4038
Db 4074 GAGGTTCAAGAGCGAGTTGAGCCGACCGTGTGATCCGCTACTAATAAGGAGAACTT 4133
Qy 4039 AGCAAGAGCGCTGACCAAGCTATCTGTAATGCTGTAATGCAAAAGTCAACAGGTTCC 4098
Db 4134 GCTATTTGTCAAGAGGAAGAGGATGTCAATGACGCAATCACTGCGACACAGAGAA 4193
Qy 4099 GGAATGCTGGTGCATCTGACCAAAATGCGCGCTCTTTGATTAACAGCCAAATGCT 4158
Db 4194 GGAATCTGCGGTGCATCTAATAAGTTGGCCGAACAGTTTCAACCGATTCAGCCACAGAG 4253
Qy 4159 GTGGGAGCGCTAGCT---TGTGAAGCAGAAACCGCTCATATCATAGCTGTAGAGCC 4215
Db 4254 ACAGTACCGGAAATGACTGTGTGTCAGAGAAAGAAAGATGATCAACGCGGTGGCCT 4313
Qy 4216 AATTTTCTAAGATCCGGAACCGGAGGCGACCTTAAGCTCGCAGCTGCTTACATGAGC 4275
Db 4314 GATTTCCGGAACACCAAGAGGAGAGCCCTGAATTTGCTGCAAAACGCTTACATGCA 4373
Qy 4276 ATAGCTTCATGTCACAGCTGAGCGGATTAACAAATATTCAGTACCGCTTACTCAAC 4335
Db 4374 GTGCAACATTAATGAATGAACATTAATATCAAGTGTGCGCATCCCACTGCTATTAACA 4433
Qy 4336 GGCATCTATTCTGTGCAAAAGATCGAGTATGCAATCATGCTACACTGTTCACTGCT 4395
Db 4434 GGCATTTACGAGCGGAAAGACCGCTTGAGGTATCACTTAATCTGTTGACAAACGCG 4453
Qy 4396 TTCGACACTAGATGCGATGTCACCATATATTTCTTGATTAACAAATGGGAGACAGG 4455
Db 4494 CTAGACAGACTGATGAGAGTAACTATCACTGCTGATTAAGAGTGAAGAAAGA 4553
Qy 4456 ATAAATGAGGCACTTCAACCGCAAGAAAGCGT---GAATTTGATGATGACAGCA 4512
Db 4554 ATCGACGCGGTGTCCAACTTAAGAGTCTGTAACTGAGTGAAGATGAGATATGAG 4613
Qy 4513 GTAGACATGACTTGTGAGGTCACCCAAACAGCTCTTTGGCAGGACAGCGATTAC 4572
Db 4614 ATCGACAGCGATTAATGATGATCAATCCGACATGCTGTAAGGAGAAAGGATTC 4673
Qy 4573 TCCGTCATGAGGCAAGTTGATTAATCACTGGAAGTACACGATTCATCAGACGCC 4632
Db 4674 AGTACTCAAAAGGAAGTTGATTTGATTTGTAAGGCAACCAATTCATCAAGAGCA 4733
Qy 4633 AAGGACATGCGGAATTCATGCAATGTGGCCCAACCAATTTGAGCTTAATGACAGATT 4692
Db 4734 AAGATATGGCGGAGATTAAGTCTGTTCCTCAATGACAGAAAGCAACGAACTG 4793

OY	4693	TGCTTGTAATCCTCGGAGGAAAGTATGTCACAGATCCGCTCCAAATGCGCCAGTAGAGAG	4752
Db	4794	TGTGCTATCATATTATGGGGAGACCATGGAACCAATCCCGGAAAATATGCCCCGTGACACC	4853
OY	4753	TCAGAGGGCTGTGCTCCACCTTCACACATTTCCATGCTGTGTAAATTACGTATGACGGT	4812
Db	4854	AACCCGTGTCTAGCCCGCCCAAAAACGCTGCGGCTCTATATGTAATGACATACGCA	4913
OY	4813	GAGCGCGTATACAGTTGTCGCTCTGCGAAGAAAGACAGTTCGCGCGTATGCTCATATTC	4872
Db	4914	GAAAGGGTCCACAACACTCAGAAAGCAATACGTCAAAAGAAAGTTACAGTATGCTCCTCACCC	4973
OY	4873	CTGTTCGCCAAGTACAGGATCAACAGGCGTGCAGAACTACAGTGAAGCAAAACCACTCTG	4932
Db	4974	CCCTCTCCAAAGTACAAATATCAAGATGTTCAGAAAGTTCAGTCCACAAAAGTATGTCCTG	5033
OY	4933	TTTTACAGCGCTGTATCCACCGGCTGTACACCCACAGAAAGTACGGGGAA	4980
Db	5034	TTTAACCGCATACCCCGCGCATTTGCTTCCGCCCGGTAAAGTACATAGAA	5081

RESULT 13

US-09-367-764-1
Sequence 1, Application US/09367764
Patent No. 6583121
GENERAL INFORMATION:
APPLICANT: Johnston, Robert E.
APPLICANT: Davis, Nancy L.
APPLICANT: Simpson, Dennis A.
TITLE OF INVENTION: System for the In Vivo Delivery and
TITLE OF INVENTION: System for the In Vivo Delivery and
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
STREET: 1211 East Morehead Street
CITY: Charlotte
STATE: No. 6583121th Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/367,764
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/801,263
FILING DATE: 19-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-147
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11663 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 60..7559
FEATURE:
NAME/KEY: CDS
LOCATION: 7608..11342
US-09-367-764-1

Query Match	26.7%;	Score 1333.2;	DB 4;	Length 11663;
Best Local Similarity	56.1%;	Pred. No. 0;		
Matches 2820;	Conservative	0;	Mismatches 2128;	Indels 80; Gaps 13

QY	23	ATATGGAAGAATTCACCGTTGACTTAGTGTGTAACAGCCCGGTATGTCAAGTGGTTACAGC	82
Dp	64	AGAACCCAGTAGTTAAAGTAGACGTAAACCCGTACAGAGTCCGTTGTTCGTGCAACTGCAAA	123
QY	83	GGACGTTTCCACAATTTTAGATCCAGACAGCAGGCTCACTGACATATGACATGCCAATG	142
Dp	124	AGAGCTTCCCAATTTAGGTAGTAGACACAGCAGGTCACTCCAAATGACCATGTCTAATG	183
QY	143	CCAGAGCGTTTCCGATGTGGCAACAAGCTCATTTGAGAGGAAAGTTCGACCCGGACCAAG	202
Dp	184	CCAGAGCAATTTTCCGATCTGGCCAGTTAACTTAATGACCTGGAGGTTCTTACACACGA	243
QY	203	TTATCTTGGACATTGGAAATGCGCCCGCTCAGACATGCATCTTCCAAATCAGCGTTATCT	262
Dp	244	CGATTTTGGACATAGGCAAGCCGACCGGCTCGTAGAATTTTTTCCGAGCACCAAGTACATT	303
QY	263	GTAATGCGCCCTATGATTAAGCGCTGAAGACCCGGACAGCTACACAGGTATGCAAGAAAGC	322
Dp	304	GCGTTTGGCCCATGTGTGATGCAGAAAGCCCGGACCGCATATGAAATATGTCACAGCAAC	363
QY	323	TTAAGAAAA-----GTGACATTACCCGACAAAGAACATAGCCCTTAAGCGGGACAGCTCAG	376
Dp	364	TGGCGGAAAAAGCATGTAAAGATTACAAACAAAGAACTTGCATGAGAAAGATCAAGAGCTTCC	423
QY	377	TGGAAGTCATGTCAACACACAGACGACAGAGACTCCATCTGTGTATGTCACACAGAGCCCA	436
Dp	424	GGACCGTACTTGATATACACCGAGTGTGAAAGCCCATCACTGCTGTTCCACAACAGATGTTA	483
QY	437	CGTGTAGTACTTTTGAAGTGTAGCAGTATACCAAGATGTGTAGCGAGTCCATGTGACCGGA	496
Dp	484	CCTGGAACACGCGTCCGAGTACTCCGTCATGACAGAGGTGTATCA--TCAACGCTCCCG	540
QY	497	CATCAATCTACACACGAGCGCTTAAAGAGTTAGGACAAATTAATCTGATAGAGCTTTGACA	556
Dp	541	GAACTATTATACACAGGCTATGAAAGCGGTGCGAACCTGTACTGGAATTTGGCTTTCGACA	600
QY	557	CGACCCCTTTTATGTACAAAAACATGGCAGGTTCTTACCTTAATTACCAACGAACTGGG	616
Dp	601	CCACCCAGTTTATGTTCTCGGCTATGGCAGGTTGTGATCCCTGCATACCAACCAACTGGG	660
QY	617	CTGACGAGAGATTTGGAAAGACAGTAACATTTGCGCTCGGTAACTCAAGTCTTTCAGAGGA	676
Dp	661	CCGACGAAAAAGTCTTGAAGCGGTAAACATGGAATCTTGTGACAGCAAAAGCTGAGTGAAG	720
QY	677	GCAAGGCTTGAANAACCTCTCAATCCTTTAGGAAGAAGGCTCCAACCTTACTAATAAGATCA	736
Dp	721	GCAAGGACGAAABATTGTCCATTAATGAGAAAGAAAGATTAAAGCCCGGGTCAAGGGTTT	780
QY	737	TATTTCTGGTTGTTCAACAATCTTACACAGAAATGATCACTGTTAAGTAGTGGCATC	796
Dp	781	AATTTCTCGTTGATGACACTTTAACCCAGAACACAGGACAGCTTGACAGAGCTGGCATC	840
QY	797	TTTCCAAAGTGTGACCTTGAAGAGAAAGTTAATTCTCACAGGTAAAGATGTGGACCAATG	856
Dp	841	TTTCATCGGTGTTCACCTTGAAGAGAAAGACGTGTAACTTGGCCGCTGTATACAGTGG	900
QY	857	TCAGCTGTGAAGGGTATAGTCAATCAAAAAAGATTAACGATCAGCCACAGACTTAATCGTTAAG	916
Dp	901	TGACTGTGAAGGCTTACGTAGTAGAAGAAATCACCATAGTCCGGGATTCAGGGAGAAA	960
QY	917	TTTGAAACTTGGGCTCAACAATGATGCGCAGGGTTTCTTGAATTGCAGAAATCAACAGTA	976
Dp	961	CCGTGGGATACGCGGTATCAAAACAATAGCGAGGCTTCTTGCTATGCAAAAGTTACCGATA	1020
QY	977	CGCTGCGGGGGAAGGGTTTCTTTTCTGTGTGTAGTATGTACAGACCAACTTTGGC	1036
Dp	1021	CAGTAAAAAGAAACGGGTATGTTTCCCGTGTGACAGTATATCCGGCCACCAATATGGC	1080

```
QY 1037 ATCAGATGACGAGGATTCTGGCAACTGACGTTAGTGTGATGACGACCAAAAACCTATTGG 1096
DB 1081 ATTCAGATGACCGGCAATATGCGCCACGATATCTCACTGACGATGACAAAACCTTCTGG 1140
QY 1097 TTGGGCTCAACCAAGATTTGCTCAATGCTAGACGCGCAAGAAATCTAACAAATGC 1156
DB 1141 TTGGGCTCAACGACGAAATGCTCATTTAACTGTAAGACTTAACAGGAAACCAATATCATGTC 1200
QY 1157 AGAATCTATCTATTACGATGCTGCGCCAGGCTTTTCCAGGTGGGCGCGTGAATCATGCTG 1216
DB 1201 AAAATTACTTCTTCCCAATCATTTGCACAAAGGTTTCAGCAAAATGGGCGCAAGAGGCGAAAG 1260
QY 1217 CCGACTTGAACGACGAGAAAGAACTAGAGGGTGGCGGAGCGCACTTACTTATGAGGCTGCT 1276
DB 1261 AAGATCTTGACATGAAAATAATGCTGGGACCAAGAGCGCAAGCTTACATATGCTGCT 1320
QY 1277 GCTGGGCTTTTCMAAGCCGAGAAATCAATCATCTACMAAGAGCTTGTAACGCAAA 1336
DB 1321 TGTGGGCTTTCCGACTAGAAAGTGCACTGTTCTATGCGCCACTGGAAACGACAGCA 1380
QY 1337 TTAAAGAAATGACTGCGCTCTTTGACTCATTTTGATTTGATTTCCAGCGCTTACAGCAAGGAC 1396
DB 1381 TCGTAAAGTCCGACCTCTTTTAGCGCTTTCCCATGTCTATCGGTATGAGCTACTCTCT 1440
QY 1397 TCGATATGGGCTTCGCGCTAGGCTCAAGCTGCTGTTGAACCA---ACTGTCAAAACCG 1453
DB 1441 TGCCATATGCTGCTAGAGGAGAAATGAAATTTGGCATTTACAAAGAGAGAGAAAC 1500
QY 1454 CACCGGCTATTACATGAGCCGATGTGAGCATCTGCGTGCTTACAGCAAGAGCTGAG 1513
DB 1501 TGCTGCAAGTCCCGAGAGAAATTAATTTAGAGGCAAGGCTGCTTGAAGATGCTCAG 1560
QY 1514 AAGTGTGCTGACGCGGAAAGATCAAGAAAGCCCTGCCACCTTGTCTCCCTAATAGAAA 1573
DB 1561 AAGATCCAGAGCGGAGAGCTCCGAGAAAGCACTCCACCATATGTCACAAAGTAA 1620
QY 1574 AAGAGCCGTAGAGGAG-----AAGTAGACCTCATTTATGCAAGAGGACAGAGAG 1624
DB 1621 TCGAGGACGCTGCGAAGTTGTCTGCAAGTGAAGGGCTCCAGCGGACACCGAGAG 1680
QY 1625 GTACGCTGAGACACACAGAGACACATCAGGCTGACAAAGTTACCCAGCGCAAGAGAA 1684
DB 1681 CACTGTGGAACCCCGCGGCTATGTAGATTAATCTCAAGCAATATGACGTATGA 1740
QY 1685 TTGGGCTTATGCTTATCTTCAACCCAGGCGGTATTAATATGTAAGAAATCTGCGTGA 1744
DB 1741 TCGGACATATATGCTGTCTGCGGATCTGTGCTGAGAAAGCGCTAACTTCGACAG 1800
QY 1745 TCCACCCATTGGCGGAAACAAATCTGTATGATGACTCAAAAGGTAGGCGAGAGATCA 1804
DB 1801 CACACCCGCTAGCAGACCAAGTTAAGATCATACGCACTCCGGAAGATCAAGAAAGTATG 1860
QY 1805 AAGTGCAGCCATACCAAGTAAAGTCAATTTGACAGAAAGGACGCGGCTCCCTGTTCAAG 1864
DB 1861 CAGTGCAGACATAGAGCTAAAGTATGATGCCAGAGAAAGTCCGTACATAGGCGAG 1920
QY 1865 ACTTCCAGGCACTTGAAGTGAAGCGGTACGATCGTTTTTAAAGAGAGGAGTTCTGAACA 1924
DB 1921 AATTTTACGACTGAGTGAAGCGCGCACTTTGTGTACAAAGAAAGAAATTTTGTGAACC 1980
QY 1925 GATACCTGACACATCGCATTCACGAGAGGCGCTAAACATGACGAGAGATCTATA 1984
DB 1981 GCAAGCTGTACATATTCATGACAGGTCGCGCTTAAGAAATTCAGAAAGAGAGAGCTACA 2040
QY 1985 AGACTGTAAAGACTCAGACACAGATCAGAAATAGCTTTCGATTTGATCCAGCAAGAAAGT 2044
DB 2041 AGGTTCACAAAGGAGAGCTCCGAGAAACAGATAGTGTGACGTGACAAAGAGCAT 2100
QY 2045 GTGTTAAGCGAAGACGAGAGTCCCTTGTGCTTAAACCGGTGATGCTGATTCACAT 2104
DB 2101 GCGTTTAAAGAGAAAGCTTCAGAGACTTGTCTTCCGAGAAATGACCAACCGGCT 2160
QY 2105 TTCAAGATTTTGCATAGAGAGTCTCAAGACACGACGACGACCTCAAAAGTCCAA 2164
DB 2161 ATACGAACTAGCTCTTGAGAGACTGAAAGCTGACCCCGGCTCCGCTACAAAGTTGAAA 2220
QY 2165 CCATCGAGCTTATGAGATCCAGGTTCAAGTAAATTTGAAATCATTAAGAAAGCTTGA 2224
DB 2221 CAATAGAGATGATAGGACACAGAGATCGGCAAGTCAAGTATCATTAAGTCAACTGTCA 2280
QY 2225 CTAAAGAAAGATCTGTGTGTAGTGCAGAAAGGAAAGAACTCGCAGAAATCATCAGGATG 2284
DB 2281 CCGCAGCTGATCTTGTTTTACAGCGGAAAGAAAGAAATCTCCCGGAAATTTGAGCGGACG 2340
QY 2285 TAAAGAGATGAGACGATAGATGTGCTGCTAGAGCTGTGATTCAGTCTCTTAATG 2344
DB 2241 TGCTAGCGCTGAGAGGAGCATGACATGACGTGAAAGACAGTGAATTCGTTATGCTCAAG 2400
QY 2345 GGGTTAAGACCCCGTTTAAACATCTGTGACATTTGATGAGGATTTGCTGCAATCAGAGGA 2404
DB 2401 GATGCCCAAGCGGTAGAGATGCTGTATGTGACGAAGGCTTCGCGTGCACGACAGAG 2460
QY 2405 CGCTGCTGGACATGATGCGATGCTCAACCC---TAAAGAAAGTGTATTTGCGGGAGC 2461
DB 2461 CACTACTTGCTGTATTTGCAATGCTCAGACCCCGTAAAGAGTATCTATGCGGAGACC 2520
QY 2462 CAAGCAATGCGGCTTCTTTTAAAGATGTGCTGAAAGTACATTTTAAAC----- 2512
DB 2521 CTAAAGCAATGCGATTTCTTCAACATATGCACTTAAAGATTAATTTCAACACCTGAAA 2580
QY 2513 ATGACATATGACTGAAGTGTACCATTAAGATCTTTAGAGGTGACACACATGTTAA 2572
DB 2581 AAGCATATGTACCAAGACATTTTCAAGTTTATCTCCAGCTGTGACACAGCAAGTCA 2640
QY 2573 CCGCATGTGCTCCAGCTCTTTAGAGCAAGGAAAGAAAGCGGTAAACCATGCTG 2632
DB 2641 CCGGTATTTGATGACACTGCTGATTTACAGTGAAGAAATGAAACCAAAACCTGTGAGA 2700
QY 2633 ATAAATCATATATGATACCAAGAGACCAAAAGCCGCAAAAGATGATCTGATTTAA 2692
DB 2701 AGAATCATGAAATGACATTAACAGGGGCAAGAGCCGAGGAGCATCATCTGTA 2760
QY 2693 CCTGTTTCAAGATGAGGTGAAGACGTACAGATTTGACTACAAATATCAAGAAATCTATGA 2752
DB 2761 CATGTTTCCGGGGTGGGTAAAGCACTGCAAAATGCACTATCCCGACATGAGGTAATGA 2820
QY 2753 CTGCGGCTGATGCAAGGACTTACGCGGAAAGGCGTTTATGCTGACAGTTAAAGTCA 2812
DB 2821 CAGCGCGGCTTCAAGAGGCTAACGAAAGAGATATGCGCTCCGCGAAAGTCA 2880
QY 2813 ACGAATTCACCTACTCTGACAGCTTCTAGACAGTGAACGTGTTACTTACAGCAG 2872
DB 2881 ATGAAACCCGCTGTACGCAATCATCAAGGATGTGAAGCTGTGCTCACCGCACTG 2940
QY 2873 AAAAAGCATTTGTGGAAGACGTACTAGTGTATCCCTGGATTAAGACACTTACACTA 2932
DB 2941 AGGACAGGCTAGTATGAAAACTTTACAGGCGGACCATGATTAACAGCTTCAACG 3000
QY 2933 AATATCCCGGGGATTTACGCGCTTATTGACGACTGCGAGCGGAGACGAGCCATTA 2992
DB 3001 TACTTAAAGAAATTTTCAAGCCACATCGAGACTGGAGAGCTGAACAAAGGAAATTA 3060
QY 2993 TGGCAGCGTTCCTTATAGCCGACAGCTGATGTGTTCCAGATTAAGGTGAAGCTCT 3052
DB 3061 TTGCTGGATTAACAGTCCGCTCCCGTTCGATTCGATTCAGTGAAGCTAACGTTT 3120
QY 3053 GCTGGGCGAAGCTTTAGAGCACTTGGCCACGCGCAACATTTGTGCTGACGACAGC 3112
DB 3121 GCTGGGCGAAGCACTGGAACCGATACCTGGCCAGCGCGGATGTGTTACCGGTTGCC 3180
QY 3113 AGTGG---GAGACCTTGCACCCATTCACAGATGACAGAGTACTCACTGAATATGACAC 3169
DB 3181 AGTGAAGAGAGCTGTCCACAGTTTGGAGTGAACCAACATCGGCAATCTACGCT 3240
QY 3170 TGAACCTCTTTGACACAGGTTCTTGGAGTAGACTGACAGTGGGTATTTTCCGCTC 3229
```

Db 3241 TAGAGTAATTTGATTAAGTTTTGGCGATGACTTGAACAGCGGCTGTTTTCCAAAC 3300
Qy 3230 CTACCGTCGCACTTACTTACAGGAT-----CAGCATGGGATA 3268
Db 3301 AGAGCATCCCTTAAGTACATCTCCGCACTCAGCGAGCCAGTACCTTGGGACA 3360
Qy 3269 ACTCCAGGGAAGAACTGTATGGGCTTAAATAGAGAGGTAGCAAAAGAGTGTACGGC 3328
Db 3361 ACAGCCAGGAACAGCAAGTATGGGTACGATCAGCGCTTCCCGCAACTCTCCGTA 3420
Qy 3329 GATATCCGTGATCAACAAAGCGGTGACACAGCGAGGTAGCTATATAGAAATA 3388
Db 3421 GATTTCCGCTGTTCCAGTACG---TGGGAAAGGACACAGCTTATTTGACAGGGGA 3477
Qy 3389 CCATCAAGGACTACTCTCCAACTTAATGTGTTTCAATTAATGGCGGTTGCCCACT 3448
Db 3478 GAACATAGATTAATCTCTGACAGCATTAATTTGTTCCAGTAAACGCAATCTCCCTCAG 3537
Qy 3449 GCTTATCGTTGACACAAAGGAGGATCAACTGATCAACGCGATTTCTATCTAAGA 3508
Db 3538 CTTATGCTCCGAGACAAAGGAAACAACCGGCCGCTGAAAATTTCTTGAGCCAGT 3597
Qy 3509 TGAAGGCAATCTGTGTGTTGATCGCGATCTTA-----TACGATTCAGAGG 3558
Db 3598 TCAACACACACTCCGTAATCTGATCTCAGAAAAAATTGMACTCCCAAGAGAA 3657
Qy 3559 AAGAAAGTATGATCATGGGTTCATTTGCCCACTAATACATAGGTGATTCGATTTG 3618
Db 3658 TCGAATGATGCCCCGATTTGGCATAGCCGGCGAGATAGAACTAACAACTGGCTTTG 3717
Qy 3619 GGAATACCTAGCATGTGCTTAATATGACATTAATCTTTGCAATGTTAGACCCCGTAC 3678
Db 3718 GGTTCGCCGAGGACG---TACGACTGGTTCATCAATATTTGAACTAAATAC 3773
Qy 3679 AGAAACATCACTACCAACAGTGCAGGATCACTGATTCACACAGCATGTTACGT 3738
Db 3774 AGAAACATCACTTCAACAGTGCAGGACACCGCGGCACTTGAACCTTTGCGGT 3833
Qy 3739 AAGGTTCACCACTGTAACCTGCGGAAACATGTGGCTATAGGGTATGGGCTTCT 3798
Db 3834 TCGGCTTGAATGCTTTAACCCCGAGGACCTGCTGGTGAATCTCTACGTTACCC 3893
Qy 3799 GATCGCGAACCGAATATCATCTGCGGTGACGCTCATTTAGGTTTACCGGTTC 3858
Db 3894 GACCCATATGAGAGAGTACTACCGCTCTGCAAAAATTTGTAGAGTGTCTGA 3953
Qy 3859 TGTCACTTAAGAACACTGCGCAAAATCTGAGGTTCTTCTGTTCTTTCGCAAGAC 3918
Db 3954 GCGAGGCCAGAGTGTCTCAAGCAATACAGAAATGTACTGATTTCCGACAACTAGAC 4013
Qy 3919 AACCGCAACACACATGACCAAGACAGCTCGGTATGTCTTGACAACTTATCA 3978
Db 4014 AACAGCCGACACGCAATTCACCCCGATCTTTGAATGTGTATTTGCTCCGTAC 4073
Qy 3979 GGGTCAACCAAGTACAGGACGAGAGAGTCCAGGCTCAGAGTATCAGAGGTGACAT 4038
Db 4074 GAGGTACAAAGAGAGAGTGTGAGCCGACCTGTGTACCTTAAGAAAGGAACTT 4133
Qy 4039 AGCAAGAGCGCTGACCAAGTATCTTATGTGTGTAATAGCAAGGTCAACAGGTTCC 4098
Db 4134 GCTGATTTCAAGAGAGAGAGTGTCTCATATGACCAATCCACTGGGACACAGAGAA 4193
Qy 4099 GAGGTGCGGTGACCTGTACCGAAATGCGCGTCTTTGATAGACAGCAATAGCT 4158
Db 4194 GAGGTGCGGTGACCTGTATTAAGGTTGGCGCAAGTTTCAACGATTCAGCCCAAG 4253
Qy 4159 GTCCGAGCGCTAGACT---TGTGAAGCAGAACCGCTCATCATCATGCTGTAGACCC 4215
Db 4254 ACAGGTACCGCAAACTGACTGTGTGCAAGAAAGAAATGATCAAGCGGTTGGCCCT 4313
Qy 4216 AATTTTCTAAGATCCGGAACCGGAGGCGCACTTAAAGCTGACGCTCTACATGAGC 4275
Db 4314 GATTTCCGAAACACCCAGAGGCGAGAGCCCTGAATTTGCTGCAAAAGCCCTACATGCA 4373

Qy 4276 ATAGGTCCATGTCACAGCTGAGCGATTAACAAATATCAGTACCGCTACTGTCAACC 4335
Db 4374 GTGGCAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4433
Qy 4336 GGCATTAATTTGTTGTCGCAAGATGAGTATGCAATCATTTGACCTGTTCTACTGCT 4395
Db 4434 GGCATTTACGAGCGGAGAAAGACCGCTTGAAGTATCACTTAATCTGTTGCAACCGCG 4493
Qy 4396 TTGCACATTAAGATGCGGATGTCACCTATATTTGTTGATTAACAATTTGGAGACCG 4455
Db 4494 CTAGACAACTGATGCGGAGTAACTCTATCTGCTGAGTAAAGAGTGAAGAGAA 4553
Qy 4456 ATTAATGAGGCACTTCAACGCAAGAAAGCGTC---GAAATTTGAGATGACAAAGCA 4512
Db 4554 ATTCAGCGGCTGCTCAACTTAAGAGTCTTAATGAGCTGAAAGATGAGATATGAG 4613
Qy 4513 GTAGCATTAATTTGTCAGGATCCACCAACAGCTTTTGGCAGGACAGCAAGTTAC 4572
Db 4614 ATTCAGCAGAGTATGATGATCCATCCGACAGTTGCTGAAGGAAAGAGGATTC 4673
Qy 4573 TCCGTCATGAGGCAAGTGTATTTATCTCTGGAAGTACAGATTTCTACAGCCGC 4632
Db 4674 AGTACTACAAAGAAAGTGTATTTGTAATTTGAAGGACACCAATTCATCAAGCAGCA 4733
Qy 4633 AAGCATTTGCCGAATCATGCAATGCGCCCAACAAATGAGGCTAATGAGCAGATT 4692
Db 4734 AAGATATGCGGAGATTAAGTCTCTGTTCCAAATGCAAGAAAGCAAGCAACATCG 4793
Qy 4693 TGCTTTATCATCTGCGGAGATATGTCAGCATCCGCTCCAAATGCGCAGTAGAGAG 4752
Db 4794 TGTGCTTACATATTTGGGAGAACATGAAACATTCGCGGAAATGCGCGGTGACAC 4853
Qy 4753 TGAAGGCTGTCTTCACTCACTCACTTCACTCTGTTATTAATGATGATGAGCGCT 4812
Db 4854 AACCGTCTGTACCGCCCAAAACGCTGCGTCTGTATGATGATGATGATGATGATG 4913
Qy 4813 GAGGCGCATACAGTTCGCTGCGGAAAGAAAGAGTTCGCGTATGCTATCATTC 4872
Db 4914 GAAAGGTCACACATCAGAAAGATTAAGTCAAGTATGATGATGATGATGATGATG 4973
Qy 4873 CTGTTGCCAAGTACAGATCAGCGCTGCGAGAGTACAGTACAGCAACAGTCTG 4932
Db 4974 CCCCTTCAAGTACAAATTAAGATTTTGAAGTGTTCAGAGTGTTCAGCAAAAGTATG 5033
Qy 4933 TTTTCAAGGCTGTACCAACGCGCTGTACACCCAGGAATGACCGGAA 4980
Db 5034 TTTAACCGCATACCCCGCATTCGTTCCCGCTGATGATGATGATGATGATGATG 5081

RESULT 14
US-09-367-764-7
; Sequence 7, Application US/09367764
; Patent No. 6583121
; GENERAL INFORMATION:
; APPLICANT: Johnston, Robert E.
; APPLICANT: Davis, Nancy L.
; TITLE OF INVENTION: System for the In Vivo Delivery and
; TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell Selitzer Park & Gibson, P.A.
; STREET: 1211 East Morehead Street
; CITY: Charlotte
; STATE: No. 583121th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

1805 AAGTCAGCCATACCGGTAAAGTCAATTGTACAGAAAGGACGGCGGTCCCTGTTCAAG 1864
1861 CAGTGGAAACATACGACCGCTAAAGTACTGATCCAGCAGAAAGTCCGATCCATGCGCAG 1920
1865 ACTTCAGGCATTTAGTGAAGCGCTAGCATCGTTTCAAGAGAGAGGATTCGTAACA 1924
1921 AATTCTTAGCCTGAGTGAAGCGCCACGCTTGTGTAACAAGAAAGAGTTTGTGAACC 1980
1925 GATTACTGACACATCCGCAATCAACGAGAGCGCTAAACAATCAAGAGTACTATA 1984
1991 GAAAGCTGACCATATGCTACATGACAGGTCCCGCTAAGATACAGAAAGAGGACGTACA 2040
1985 AGACTGTAAGAGCTCAGACAGACAGACTCAGAAATACGTTTCGATATTGACGACGAAAGT 2044
2041 AGGTTACAAAGGACAGAGCTCGAGAAACAGAGTACTGTTGACGTGACAAAGAAAGCAT 2100
2045 GTGTTAAGGAGAGACGCGAGGTCCCTGTCCTAAACCGGTATCTGTAGATCCACAT 2104
2101 GCGTTAAGAGAGAGAGGCTCAGGACTTGTCTTTCGAGAGATGACCAACCCGCCCT 2160
2105 TTCAGAGTTTCCGTAACGAGAGTCTTCAGACAGACAGACACTTCACAAAGTCCAA 2164
2161 ATCAGGAACCTGCTTGAAGGACGTGAAGACTGACCCGCGTCCGTAAGAGTTGAAA 2220
2165 CCATCGAGTCTATGAGAGTCCAGGTTCAAGTAAATCTGATATCAAAAAGCGCTGTA 2224
2221 CAATTAGAGATATAGGCACACAGAGTCCGAGACAGTCACTATCAATGATCAATCTTCA 2280
2225 CTAAAGAAAGATCTGTTGTGAGTGCAGAGAGAGAAACCTGCGAGAAATCATCAGGATG 2284
2281 CGGCACTGATCTTGTACACGCGAAAGAAAGAAATGCGCGAAATTAAGGCGCAGC 2340
2285 TAAGAGATGAGAGATGATGATGTTGCTGCTAGAGACTGTCATTCAGTCTTCTTAATG 2344
2341 TGTACGCGTGAAGGACATGAGATCACTCGAAGACAGTGAATTCGTTAATGCTCAACG 2400
2345 GGGTTAAGCAACCCGTTAAACACTGTATCATTTGATGAGGCAATTCGCTGCCATGACGGA 2404
2401 GATGCCAAGAACCCGTAGAAAGTCTGTATGTTGAGAGACGTTCCGGTCCACGAGAG 2460
2405 CGCTGTCGACATGATGTCATCTGTAACCC--TAAGAAAGTGTATTTGTCGGGAGC 2461
2461 CACTACTGCTTGAATTTGCAATGTCAGACCCCGTAAGAGATGATGATGCGAGAGC 2520
2462 CAAAACATGCGGCTTTTAAATGATGTCCTGAAAGTACATTTAAAC----- 2512
2521 CTAAACATGCGGATTTCTTCAACATGATGCACTAAAGGTACATTTCAACCCCTGAAA 2580
2513 ATGACATATGCACTGAAGTACCAATTAAGCATCTCTAGAGGTGCAACAGACTGTAA 2572
2581 AAGACATATGTAACAAGACATCTACAAATTTATCTCCGACGTGCAACAGCAGTCA 2640
2573 CCGCATCGTCTCAACGCTCTTCTAGCAACAGATGAAAGAGGTTAACCATGTGCTG 2632
2641 GGGCATATGTAACGACATGATTAAGTGAAGAAATGAAGAACCAAAACCCGTGCAAGA 2700
2633 ATAAATATCATATGATACACAGGAGACCAAAAGCCGCAAAAGATGATCTGATTTAA 2692
2701 AAGAAATGAAATTCACATTAAGGAGGACCAAGAACCCGAGCAGATCATCTGTA 2760
2693 CCTGTTTCAAGAGATGAGTGAACAGCTACAGATTGACTACAAATAATACGAAATCAAGA 2752
2761 CATGTTTCGCGGGGTGAAGCACTGCAAAATGACTATCCGAGCATGAGGTAAGA 2820
2753 CTGCGGCTGATCGCAAGACTTACCGGAAAGGCGTTTATGCTGTCAAGGTCAAAAGTCA 2812
2821 CAGCGCGGCTTCAACAGGCTTAACAGAAAGGATATGCGCTCGGCGGCAAAAGTCA 2880
2813 ACGAATTCATCTACTGCAAGCTTCTGAGACAGTGAACGTGTTACTTACAGCAGAC 2872
2881 ATGAAACCCGCTGTACGAGTACATGACAGCATGTGAACGTGTTGCTACACCCGACAG 2940

2873 AAAACGATTTGCTGGAAGACGCTAGCTGTGATCCCTGATTAAGACACTTACAGCTA 2932
2941 AGGACAGGCTATGATGAAGAACTTTACAGGGGAGACCATGATTAAGCACTCATTAACG 3000
2933 AATATCCGGGAGATTTACGGCTTCATTTGACAGACTGACAGCGGACGACAGCCATTA 2992
3001 TACCTAAAGAAATTTTACAGGACCATGAGAGATGCGGAAGCTGAACAAGAGAAATTA 3060
2993 TGGCAGCGGTTCTGTATTAAGCCGACAGACGCTGATGTTTCCAGAAATTAAGGAAAGCT 3052
3061 TTGCTGCATTAACAGTCCCGCTCCCGTACCAATCCGTTACGCTCAAGACTTAACGTTT 3120
3053 GCTGGCGAAGGCTTTAAGCAGAGTCTTGCCCAACGCGCAACATTTGCTGACAGACAGC 3112
3121 GCTGGCGAAGGACATGGAACCGATACCTGCGCACGCGCGGTATGCTACTTACCGTTGCC 3180
3113 AGTGG--GAGACGTTGACCCATTCAGATGACAGAGCGTACTCACCTGAATGCGAC 3169
3181 AGTGAAGGAGCTGTTCCACAGATTGCGGATGACAAACACACACTCGGCAATCTACGCT 3240
3170 TGAATCTTTTGGACAGGTTCTTTGAGATGAGACTGGAACAGTGGTTATTTCCGCTC 3229
3241 TAGAGTAATTTGATTAAGTTTTCGGCTGAGACTGACAAAGCGGCTGTTTTCAAAC 3300
3230 CTACCGTGCATTACTTACAGGAT-----CAGCACTGGAGTA 3268
3301 AGAGATCCCGTTAAGTACATCATCTGCGCATACGACAGCCAGTACTGTTGGACA 3360
3269 ACTGCGAGGAGAAAGATGATGAGCTTAAATGAGAGTGAACAAAGAGTTGTCACGCG 3328
3361 ACAGCCAGAAACACGCAATGATGGGTACGATACAGCGGTTGCGCGCAACTCTCCGTA 3420
3329 GATATCCGTGATCACAAAACCGTTGACAGAGAGGATGATGATTAAGATTAATA 3388
3421 GATTTCTGCTGTTCAAGTAC--TGGAAAGGACACAGCTGATTTGACAGAGGCA 3477
3389 CCATCAAGAGACTACTCTCCAAATTAATGATGTTTCAATTAATGCGCGGTTGCCCACT 3448
3478 GAATAGATATCTCTGACAGATTAATCTGTTCCAGTGAACCGCAATCTCTCCACG 3537
3449 GCTGATGTTGACCAACAAAGACAGGTTACATGATCAAGCGGATTCATCTAAGA 3508
3538 CTTATGTCGCCGAGACAAAGAGAAACACCCGCGCGGTGAAAATTTCTTGAACGAGT 3597
3509 TGAAGGCAATCTGTGTTGATGATCGGATCTTA-----TAGCATTCACAGG 3558
3598 TCAACACACCTCGTACTTGTGATCTCAGAGAAATAATGAAGCTCCCAACAGAGAA 3657
3559 AAGAAAGTATGATCATGGGTCAATGGTCCACTAATACATCAAGTGTGATCTGATTTG 3618
3658 TCGAATGATGCGCCGATGTCAGTACGCGGACGAGATTAAGAACTAACCTGCTTTCG 3717
3619 GAAATACCTTACCATGTCGTAAATATGACATTAATCTTTGTCAATGTTAGACCCCGTAC 3678
3718 GGTTCGCGCGCAGGACAGG----TAGACCTGAGTTTCAATATTAATGAACTAAATAC 3773
3679 AAGAACTATCTACCAACAGTGGAGGATACGCTATCAACACAGCATGCTAACGTGT 3738
3774 AAGAACTATCTTCAACAGTGGAGAACCAACGCGGAGCACTTGAATAACCTTTGCGGT 3833
3739 AAGGCTGTCACACCTGAAACACTGGCGGGAACATGTGGGCTATAGGCTATGAGGCTTCT 3798
3834 TCGGCTCTGAATCTGCTTAACCCGAGAGAGACCTCGTGTGAGAGTCTTACGCTTACG 3893
3799 GATGCGCAACCGAATATCATCTGCGGTGACGCTCATTTAGGTTTAAACCGGTTC 3858
3894 GACCGCAATTAAGAGAGTATGATCAACGCTCTTTCAGAAATTTTGAAGTGTCTCA 3953
3859 TGTACGCTTAAGAACATCTGCGGAAATTAAGAGTTCTTCTGCTGTTTTCGCGAGAGAC 3918
3954 GCGAGGCGAGTGTGCTTCAAGCAATACAAATTAATCTGATTTTCCGAACTACAGAC 4013
3919 AAGCGCAACGACACATGACAGAGACAGACTCGGTGATGCTTGAACATCATATCA 3978

```

Db      4014 AACGCCGACACGACCAATTCACCCCGCATTTGAATGTGTGATTTGTCCTGATAC 4073
Qy      3979 GGGTCAACAGAGTACGAGGACGAGAGCTCCAGCTTACAGAGTACAGGTGACATT 4038
Db      4074 GAGGGTACAAAGACGAGGTGGAGCCGACCGTGTACCTTACTTAAAGGAGAACATT 4133
Qy      4039 AGCAAGACCGGTGACCAAGCTATGCTTATGCTGTATATAGCAAGGTCACCAAGTTCC 4098
Db      4134 GCTGATGTCTAAGAGGAGAGGAGTGTCTAATCAGCCCACTGGGAGCCAGAGAA 4193
Qy      4099 GAGATGTGCGGTGACAGTGTACCAAAATGCGCGCTGCTTTGATAGACACCAATAGCT 4158
Db      4194 GAGAGTGTGCGGTGACATCTATTAAGTGTGGCCGAAGATTTCACCGATTTCAGCAGAG 4253
Qy      4159 GTCCGAGCGGTAACT---TGTGAAGACGACCGCTCATCATCATCTGTGTAGACCC 4215
Db      4254 ACAGGTACCGGAAATGACTGTGTGCCAAGAAAGAAATGATCCAGCGGTGGCCCT 4313
Qy      4216 AATTTTCTAAGATGCCGGAACCGAGGCGGACCTTAAAGTCCGAGCTTACATAGAGC 4275
Db      4314 GATTTCGGGAAACACCCAGAGGACAGAGCCCTGAATTTGCTGCAAAAGCCTTACATGCA 4373
Qy      4276 ATAGCGTTCATGTCACAGCTGAGCGGATTAACAAATATGATACCGCTACTGTCAAC 4335
Db      4374 GTGCGAGACTTAATGAATGAACATTAATCAAGTGTGTGCCATGCCACTGTCTATTA 4433
Qy      4336 GGCATCTATTTGTTGGTGGAAGATGAGATGATGATGATGATGATGATGATGATGATG 4395
Db      4434 GGCATTTAAGCAGCGCGGAAAGACCGCTTGAAGTATCTAATCTGTTTGAACACCGCG 4493
Qy      4396 TTCACACTACGATGCGGATGTCACCATATATATGCTTGAATTAACAATGGAGACCGAG 4455
Db      4494 CTAGACAGAACTGATGCGGAGTGAATCAATCTAATCTGCTGATTAAGAGTGAAGAAAG 4553
Qy      4456 ATAAATGAGGCCATTTCACCGGAAAGAAAGCGTC---GAAATTCGATGATGACAAAGCA 4512
Db      4554 ATGACGCGGTGCTCAACTTAAGAGTCTGTAATCTGATGAGATGAGATATGAG 4613
Qy      4513 GTACACTGATGCTTGGTCAAGGTCACCCAAACAGCTTTTGGCAGGACGACAGGTTAC 4572
Db      4614 ATGACGACGAGTATGATGATCCATCCGAGAGTTCCTGAAGGAAAGGAGATTC 4673
Qy      4573 TCCCTCATGAGGCGCAAGTGTATTCATACCTGGAAGGTACACGATTCATCAGACCGCC 4632
Db      4674 AGTACTACAAAGAAAGTGTATTTCTGATTTGAAGGACCAATTTCCATCAGACGCA 4733
Qy      4633 AAGGACATTTGCCGAATTCATGCAATGTGCCCAACAAATCTGAGGCTAATGACAGATT 4692
Db      4734 AAGATATGCGGAGATTAAGTCTGTTCCAAATGACAGAAAGCAAGCAACTG 4793
Qy      4693 TGCCTTATCATCTCTGGGAGAGATGTGTCAGCATCCGCTCCAAATGCCAGTGAAGAG 4752
Db      4794 TGTCTCATCAATATTTGGGGAGACCATGGAACCAATCCCGGAAATATGCCCGGTGACAC 4853
Qy      4753 TCAGAGGCGTGTGCTCCACCTCACACACTTCATGCGCTGTGTAATTAAGCTATGAGCGCT 4812
Db      4854 AACCCTGTGTAGCCCGCCAAAGAGCTGCGTGTCTGTATGTATGCTATGAGGCA 4913
Qy      4813 GAGCGCGTATACAGGTGCGCTGTGCGAAGAAAGACAGTTCCGCTATGCTCATATTC 4872
Db      4914 GAAAGGGTCCACAGACTCAGAAAGCAATACGTAAGAAAGTTACGATGTCTCCACAC 4973
Qy      4873 CTGTGCGGAAGTACAGATCAGAGGCGGTGAGAAAGCTACGTCAGCAAGACAGTCTG 4932
Db      4974 CCCCTTCCAAAGTACAAATATGAATGTTCAGAAAGTTCAAGTCACAAAAGTATGCTCG 5033
Qy      4933 TTTTACGCGGTGACACCGGCTGTACACCCGAGAAAGTACGGGAA 4980
Db      5034 TTTTACCGGATACCCCGCATTTGTTCCGCGCTGATAGTATAGAA 5081

```

```

US-08-801-263A-8
: Sequence 8, Application US/08801263A
: Patent No. 5811407
:
: GENERAL INFORMATION:
: APPLICANT: Johnston, Robert E.
: APPLICANT: Davis, Nancy L.
: APPLICANT: Simpson, Dennis A.
: TITLE OF INVENTION: System for the In Vivo Delivery and
: NUMBER OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
: CORRESPONDENCE ADDRESS:
: ADDRESS: Bell Seltzer Park & Gibson, P.A.
: STREET: 1211 East Morehead Street
: CITY: Charlotte
: STATE: No. 5811407ch Carolina
: COUNTRY: USA
: ZIP: 28234
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/801,263A
: FILING DATE: 19-FEB-1997
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Sibley, Kenneth D.
: REGISTRATION NUMBER: 31,665
: REFERENCE/DOCKET NUMBER: 5470-147
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 919-420-2200
: TELEFAX: 919-881-3175
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 11703 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-08-801-263A-8
:
: Query Match 26.1%; Score 1303.6; DB 1; Length 11703;
: Best Local Similarity 56.1%; Pred. No. 0;
: Matches 2821; Conservative 0; Mismatches 2129; Indels 80; Gaps 16;
:
: 23 ATATGAAAGAAATTCACGTTGACTTGAATGCTGACAGCCCGTATGTCAGTCTTACAGC 82
: 64 AGAAGCCAGTAGTAAGTAGAGTAGAGTACAGACGATCCTCCAAATGACCAATGCTATG 183
: 83 GAGCGTTTCCAAATTTGAGATGAGAGGACGAGGTCACCTGCAATGACCATGSCAATG 142
: 124 AAGCTTCCGCAATTTGAGATGAGTACAGACAGAGTCACTCCAAATGACCAATGCTATG 183
: 143 CCAAGCGTTTTCGATGTGGCAACAAAGCTCAATTGAGAGCGAAGTGCAGCGGAGCCAG 202
: 184 CCAAGCATTTTCCGATCTGGCAGTAATACTAATCGAGCGAGGTTCTTACCAACAGCA 243
: 203 TTATCTTGAACATTTGAAGTGGCGCGCTGACAGATGCAATTCATTCACCGCTATCAT 262
: 244 CGATCTTGAACATTAAGGAGCGACCGGCTGTGATGATGTTTTCGACACAGATATATT 303
: 263 GTATCTGCGCTTATGATTAAGGCTGAAGACCGGAGCAGACTTCAACAAAGTATGCAAGAA 322
: 304 GTGTCTCCCATAGCTGATGTCAGAAAGACCGGAGCGCATGATGAATATATGCAAGTAA 363
: 323 TTAAGAA-----GTGACATTACCGCAAGAACATAGCCTTAAAGCGGAGACCTGC 376
: 364 TGGCGGAAAGAGGTGGAATTAACAAGAACTTGATGAGAAATTAAGATCTCC 423
: 377 TGAAGTCATGTCAACACAGAGCAGAGACTCCATCTTGTGTATGACACAGACCGCA 436
: 424 GAGCGTACTTGAATACCGCGAGTGTGAACACCATGCTCTGCTTCAACAGATGTTA 483

```

437 CGTGTAGTACTTGGAAAGTGTAGCATATACAGATGTGTACGACGTCCATGCACGA 496
484 COTGAACATGCGTCCCAATATTCCTCATGTAGACGCTGTA---TATCAACGCTCCG 540
497 CATCATCTTACCACGAGCGCTTAAAGAGTTAGACAAATTTATGTGATGGCTTTGACA 556
541 GAACTATCTATCATGAGGCTATGAAAGGCGTCCGACCTGTATGATTTGCTTCGACA 600
557 CGACCCCTTTTATGTACAAAACATGGCAGGTTCTCACTTACTTACAAACAGAACTGGG 616
601 CCACCACTGATGTTCTGCGCTATGCGAGGTTCTGATCCCTGCTGTAACAACCAACTGGG 660
617 CTGACGAGAGATATTGGACAGCTTAACATTTGGCCTCGGTAACTCAATCTTCAAGAGA 676
661 CCGAGGAAAGATCTTTGAAAGCGCTTACATCGACTTTGACGACAAAGCTGATGAAG 720
677 GCAGGCTTGAAAACCTTCAATCTTAAAGAGAGAGGCTCCAACTTAAATGAATCA 736
721 GTAGGACAGAAAATTTGATATATGAGGAAAGAGTTGAAGCCCGGCTCGCGGTTT 780
737 TATTCGTGTGTTCATCAATCTACAGAAAGATATGATCACTGTATGATGCTGGCATC 796
781 ATTTCTCGTAGGATCGACACTTATCCAGAACAGAGCCAGCTTGACAGCTGCAATC 840
797 TTCCAACGCTTCCACTTGAAGAGAAAGTCTTAACTTCAAGTATGTGGGACCATTC 856
841 TTCCATCGGTGTTCATCTTGAATGAAAGAGCTGTAACCTTCCGCTGTGATACATG 900
857 TCAGCTGTGAAGGATGATCATCAAAAGATATACGATCAGCCAGACATATACGTTAAG 916
901 TGAGTTGGAGAGGCTACGATGAGAAATCAACATCATGCTCCGGGATCAAGGAGAAA 960
917 TTGAAACTTGGCGCTCCACATGATCCGAGAGGTTTCTTGAATTGCAAAAGTCAAGATA 976
961 CCGTGGGATACGCGGTTACACAAATAGAGAGGCTTCTTGTATGAAAGTTACTGACA 1020
977 CGCTCGCGGAGAGAGGTTTCTTGTGTGTATGATATGATATACAGCAACATTTGG 1036
1021 CAGTAAAGAGAAAGGATATGCTTCTGTGTGACATGATCCCGGACCAATATGCG 1080
1037 ATCAGATGACAGGATTTCTGGCACTGACGTTAGTGTGATGACGACAAAAATATTGG 1096
1081 ATCAGATGATGATATATGCGCACGATATATCACTGACGATGACAAAAATCTTCCG 1140
1097 TTGGGCTCAACCAAGATGTGTGATGATGATGATGATGATGATGATGATGATGATG 1156
1141 TTGGGCTCAACCAAGATGTGTGATGATGATGATGATGATGATGATGATGATGATG 1200
1157 AGAATATCTATTTACAGTGTGCGCCAGGCTTTTCAAGGTGGGCGGTGACATCTGTG 1216
1201 AAAATTTACTTCTGCGATCATAGACAAAGGTTCAAGAAATGGGCTAAAGGCCAAG 1260
1217 CCGACTTGGACGACAGAAAGAACTAGGAGTGGCGGAGCGCACTTTATATGAGCTGCT 1276
1261 ATGATCTTGATTAAGAGAAAATGCTGGGTATCTAGAGAAAGCAAGTTTATGATGCTGCT 1320
1277 GCTGGGCTTTTCAAGACCAAGAAATCAATCATTTACAGAAAGCTGTGACGAAACA 1336
1321 TGTGGGCTTTTCAAGACCAAGAAATCAATCATTTTATGCGCCCACTGGAAGCGACACA 1380
1337 TTAAGAAAGTACCTGCGCTTTTGAATGATGATGATGATGATGATGATGATGATGATG 1396
1381 TGTAAAGATCCCAAGCTCTTTTATGAGCTTTTCCATGCTGCTGATGAGAGACCTCTT 1440
1397 TGTATATGAGCTTCCGCGTATGCTCAAGCTGCTGCTTGAACCA---ACTGTCAAAACCG 1453
1441 TGCCATGATGCTGAGGAGAAATTTGAATCTGGCATTTGCAACAAAGAGGAGAAAC 1500
1454 CACCGGCTATTTCAATGCGCATGTGAGCATCTGCGTGGCTTACAGCAAGAGCTGAAG 1513
1501 TGCTGAGGCTCTCGAGGAAATTTATGATGAGAGCCAAAGGCTGCTTTTGAAGATGCTCAG 1560

1514 AAGTGTGACGCGGAAAGATCAGAGAGCCCTGCCACCTTGCTCCCTGAATAAGAAA 1573
1561 AGGAACCGAGCGGAGAAAGCTCCGAGAAAGCATTTCAACATTTATGTGACGAAAGCA 1620
1574 AAGAGCCGTAGAGCA-----GAATGACCTCATTTATGCAAGAGCGAGAGCAG 1624
1621 TCGAGGCGAGCGGAGAAATGTTGTCTGCGAAGTGAAGGAGGCTCCAGCGGACATCGAGCGAG 1680
1625 GTAAGCTGAGAACCAAGAGAGACATCAGAGTGACAAAGTTAACCCAGCGGAGAGAGA 1684
1681 CATTAAGTTAAACCCCGCGGTACGTAAGGATATATCTTCAAGCAAAATGACCGTATGA 1740
1685 TTGGGTCTTACGCTATCTTCAACCCAGCGGATTAATATGATGAAAAAATCGGCGGTGA 1744
1741 TCGACATATATGCTGTTCTGCGCAACTCTGTCTGAAAGATGCCAACTCGACACAG 1800
1745 TCCACCAATGGCGGAAACAGTACTGTAATGATCTACAAAGATGAGGCGAGAGATCA 1804
1801 CGCACCCGCTAGCAGATCAGGTTAAAGATCATTAACACATCCGCTGATCAGGAAAGTAC 1860
1805 AAGTGCAGCCATPACCGAGGTAAGTCAATTGTAACGAGAGGAGCGGCTCCCTGTTCAAG 1864
1861 CGGTGAAACCATACGAGCTTAAAGTATGATGCCAGAGAGGTCCGTATCATGCGCAG 1920
1865 ACTTCCAGGCAATGAGTGAAGAGCGCTACGATCGTTTCAACGAGAGAGGATCGTAAACA 1924
1921 AATTCCTAGCACTGATGATGAGAGCGCACGTTATGATTAACGAAAGAGATTTGTAAC 1980
1925 GATACCTGACACCATCGCAATCAACGAGAGAGCGTAAACATGACCAAGAGTACTATA 1984
1981 GCAAACTATACCATCTTCCATGATGAGCGCCCGCAAGAAATACAGAAAGAGGAGCTACA 2040
1985 AGACTGTAAAGATCAGACACAGACTAGTAATAGTCTTGATTTGAACGACGAAAGT 2044
2041 AGTTTAAAGAGAGAGGCTTCAAGAAACAGATGATGATTTGAACGTTGACAAAGAGCT 2100
2045 GTGTTAAGCGAAGAGACGCAAGTCCCTGTGCTTAAACGCGTATGATGATGATGATGAT 2104
2101 GCGTTAAGAGAGAGAGGCTCAGGTCTGTGCTTCTGCGGAGAAATGACCAACCTCTCT 2160
2105 TTCAAGATTTGCGTACGAGAGTCTCAAGACAGACGACGACCTGACAAAGTCCCA 2164
2161 ATCATGAGACTGCTGTGAGGAGACTGAAGACCCGACCTGCGGTCCGTAACAAAGTCA 2220
2165 CCATGCGAGTCTATGAGAGTCCAGGTTCAAGTAAATCTGGAATCATCAAAAGCGTGA 2224
2221 CAATAGGATATAGGACACCGGCGTGGCAAGTCAATTAATCAAGTCAATGTCAC 2280
2225 CTAAGAAAGATCTGTTGTGATGAGTGAAGAGAAAGAACTGGCAGAAATCATCAGGATG 2284
2281 CCGACGCGGATCTTGTTCACGCGGAAAGAAAGAAATTTGCGCAAAATGAGCGGACG 2340
2285 TAAAGAGATGAGAGATATGATGTTGCTGCTAGAGACTGTGATTCAGTGTCTTAAATG 2344
2341 TGTAAAGATGAGGAGTATGAGATTAAGTGAAGAGACGATGATTCGTTATGCTCAAG 2400
2345 GGGTTAAGCAACCGGTTAACACTGTATCAATTTAGAGGCAATTTGCTGCCATGAGGCA 2404
2401 GATGCCAAGAGCGTGAAGATGCTGTATGTCGAAAGCTTCCGTCGACAGAGAGAG 2460
2405 CGCTGCTGGCACTGATGCGATGCGTCAAAAC---TAAAGAAAGTGTATTTGGCGGAGAC 2461
2461 CACTACTTCTGATGATGCTATGCTGACGAGCCCGCAAGAGGATGATCTATGCGAGAC 2520
2462 CAATAATGCGGCTTTTAAACATGATGCTGTAAGATCAATTTTAAAC----- 2512
2521 CCAATGCAATGCGGATTTCTTAAACATGATGCAATTAAGTCAATTTCAATCAACCTGAAA 2580
2513 ATGACATATGCACTGAGATGATCAATTAAGATCTTGAAGAGTGCACACAGACTGTA 2572
2581 AAGACATATGCAACAGCATTTCTACAAATATCTCCGCGGCTTGCACACAGCGATTA 2640
2573 CCGCATGCTCTCAACGCTTTTACGACAAAGCAAGTAAAGGTTAAACCATGATGCTG 2632

Db 2641 CAGTATTTGATTCGACACTGTCATTACGATGGAAAGATGAAACCAACGAAACCCGGCAAGA 2700
Qy 2633 ATAAATATCATATGATATACCAAGAGGACCAAAAGCCGCAAAAGATGATCTGATTCTAA 2632
Db 2701 AGAACATTTGAAATGATATTAACAGGGGCGCAAAAGCCGAAAGCCAGGGATATCATCTCGA 2760
Qy 2693 CCTGTTTCAGAGATGGGTGAAACAGCTACAGATTGACTACAAATAATCAGAAATTCATGA 2752
Db 2761 CATGTTTCGGGGGTGGGTAAAGCAATTGCAATTCGACTATCCGGACATGAAGTAATGA 2820
Qy 2753 CTGGGGCTGATCGCAAGAGCTTAACGGGAAAGCGTTTATGCTGTCAAGTACAAAGTCA 2812
Db 2821 CAGCGCGGCTTCACAAAGGCTTAACAGAAAGAGGTGTATCCGTCGGCAAAAGTCA 2880
Qy 2813 ACGGAATCCACTCTACTTCGACAGCTTGTGACAGTGAAGTGTACTTACACGCACAG 2872
Db 2881 ATGAAACCCCATGTGACCGCATCATCAGACATGTGAACGTGTGCTCACCCGCACTG 2940
Qy 2873 AAAAAGCATTTGTGGAAGACGCTAGCTGTGATCCCTGGATTAAGA CATTACAGCTA 2932
Db 2941 AGGACAGGCTAGTGTGAAAACTTGCAGGGCGACCATGATTAAGACGCTCACTAACA 3000
Qy 2933 AATATCCGGGGATTTTACGGGCTTCATTGACGACTGGACGCGGAACACGACGCCATTA 2992
Db 3001 TACTTAAGGAACTTTACAGGCTACTATAGAGAGCTGGGAAGCTGAACACAAAGGAATTA 3060
Qy 2993 TGGACCGCTTCTTGATTAAGCCGACAGACGCTGATGTGTTCCAAATTAAGTGAACGCT 3052
Db 3061 TTGTCGAATTAACAGCCCACTCCCGTGCATTCGTTCACTGCAAGCAACAGCTT 3120
Qy 3053 GCTGGGCGAAGGCTTTAGAGCCAGTCTTGGCCACGCGCAACATTGTGCTGACGACAGC 3112
Db 3121 GCTGGGCGAAGCATTTGGAACCGATCTAGCCACGCGCGGTATGTACTTACCGGTGCC 3180
Qy 3113 AGTGG---GAGACGTTGCAACCCATTCAGACATGACAGAGCGTACTCACCTGAATGCGAC 3169
Db 3181 AGTGGACGAACTGTTCCACAGTTTGGGATGACAAACCAATTCGCGCATTTACGCGCT 3240
Qy 3170 TGAATCTTTTGAACCAAGTCTTTGAGTGAAGACCTGGACAGTGGGTATTTCCGCTC 3229
Db 3241 TAGACGTAATTTGATTAAGTTTTCGGCATGACCTTGACAAAGCGGACTGTTTCTAAAC 3300
Qy 3230 CTACCGTCGCACTTACTTACAG-----GGATCAGACACTGGGATA 3268
Db 3301 AGACATATCCCATTAACGATACATCCCGCCGATTCAGACAGCCGCTAGCTATTTGGACA 3360
Qy 3269 ACTGCGCAGGAGAAACATGTATGGGCTTAATTAAGAGATGACAAAGAGTGTTCACGCG 3328
Db 3361 ACAGCCCAAGAAACCCGCAAGTATGGGTACGATCAGCATTCGCGCGAATCTCCGTA 3420
Qy 3329 GATATCCGTGATCACAAAACCGGTTGACACAGCAAGGTAGCTGATTAAGAAATATA 3388
Db 3421 GATTTCCGGTGTTCAGCTAC---TGGGAAGGGACACAACTTGATTGACAGCGGGA 3477
Qy 3389 CCATCAAGAACTACTCTTCAACAATTAATGTGTTCATTAATGCGCGGTTGCCCACT 3448
Db 3478 GAACCAAGATTAATCTTGCAACAGATTAACGTGTCGCGGTAAACCGCAATCTTCTCAG 3537
Qy 3449 CGTTGATGTTGACCAAAAGGATGACAACTGATCAGCGGATTCCTATCTAAGA 3508
Db 3538 CTTATGTCGCCGATACAAAGAGAACAAACCGCGCGGTGAAATAATTTTGAACAGT 3597
Qy 3509 TGAAGGCGAATCTGTGTGTGTATC---GGCGATCTATCAGCAATTCAGAGGAAGAA 3565
Db 3598 TCAAAACCACTCATGTAATCTTGATATCAGAGAAAAAATGAAGCTCCCGCTAAGAA 3657
Qy 3566 TAGAGTCATGGTTCATTTGCCAC-----TAATACATCAGGTGATCTCGATTTGG 3619
Db 3558 TCGATGATGCGCCCGCATTTGGCATAGCCGCTCAGATTAAGAACTACACCTGGCTTTCG 3717
Qy 3620 GAATACCTAGCCATGTGGTAAATATGACATTAATCTTTGTCAATGTTAAGACCCCGTACA 3679

Db 3718 GTTTTCGCGCAGGACAG---GTACGACCTGTGTTTCAATCAATTTGAACATAATACA 3774
Qy 3680 GGAACCATCACTAACCAAGGAGGATCAGCTATCCACACAGCATGTAACTGTGA 3729
Db 3775 GAACACCACTTTACACAGTGTGGAAGCAATGCGGACCTTAAAAACCTTTCCGCTT 3834
Qy 3740 AGGCTGTCAACCACTTGAAACCTGGCGAAACATGTGTGCTAATAGGATATGGCTTCTG 3799
Db 3835 CGGCTCGAATTTGCTTAACCCAGAGGACACCTCGGTGTGAAGTCTTATGGCTAACCCG 3894
Qy 3800 ATGCGCAACCGGAATATATCATCTCGCGGTGACACGCTATTAAGTTTATCCGTGCT 3859
Db 3895 ACCGCAACAGTGAAGAGTGTACACCGCTTTGCAAAAGATTGTCCAGGCTGCCAG 3954
Qy 3860 GTACGCTTAAGAACACTGCCGAAATATCTAGGTTCTTTCGTGTTCTTCCGCAAGACA 3919
Db 3955 CGAGCCAGATTTGTGTTCAAGCAATACAAATATGATCTGATTTTCCGACACTAAGCA 4014
Qy 3920 ACGGCAACACACATGACACAGACAGACTCGGTGATGCTTGAACAATCTAT-CAA 3978
Db 4015 ACAGC-GTACAGGCAATTCACCCGCACTCATGTAATTCGATTTGCTCGCTGTAT 4073
Qy 3979 GGGTCAACCAAGTACGAGGCAAGGAGCTTCAGCTTACAGATGATCAAGTGAAT 4038
Db 4074 GAGGTTCAAGAGATGAGTGGAGCGCGCCGCTCATACCGCACAAAGGAGGAATATT 4133
Qy 4039 AGCAAGAGCGCTGACCAAGCTATGTTATGTGCTAATAGCAAAAGTCAACAGCTTCC 4098
Db 4134 GCTGACTGTCAAGAGAGAGAGCTTTTCAAGCCAGCAATCCGCTGGTAAACAGCGGA 4193
Qy 4099 GGAATGTGCGGTGACCTGTACCGAAATGCGCGCTGCTTTTGAATAGACAGCAATAGCT 4158
Db 4194 GGAATGTGCGGTGACCTGTATTAACGTTGCGGACGATTTTACCAATGACCCAGAG 4253
Qy 4159 GTCGGACGCTAGACT---TGTGAACAGAAACGCTCATATCATCTGTATGACACC 4215
Db 4254 ACAGGCAACCCCAAGAAAGTACTGTGTCTAGAAAGAAAGTATCCAGCGGTGGCTT 4313
Qy 4216 AATTTTCTAAGATGCGGAAACCGGAGGCGACCTTAAGCTCGAGCTGCTACATAGC 4275
Db 4314 GATTTCCGAAACACCCAGAAAGAGAGCTTGAATTTGCTACAAACGCTTACATGCA 4373
Qy 4276 ATAGCGTCACTGTCACGCTGACGCGATTACAAATAATCAATACCGTACTGTCAAC 4335
Db 4374 GTGCGACATTAATTAATGAACATTAATCATCAAGTCTGTCCATTCAGCTGATATACA 4433
Qy 4336 GGCATTAATCTGTGCAAAAGATCGAGTATGATCAATTTGATCTCTTTTCACTGCT 4395
Db 4434 GGCATTTACGACGCGGAAAGAACCGCTTGAAGTATCACTTAACTGTGAACAACGCG 4493
Qy 4396 TTGACACTACGAGTGCAGATGTCAACATTAATTTGATTAACCAATGGGAGACAGG 4455
Db 4494 CTAGACGAATCTAGCGGACGTAACATCTATTGCTGTGATTAAGAGTGAAGGAAGA 4553
Qy 4456 ATATCAGAGCACTTCAACGCAAAAGAAAGCGTGAATTTCT---GGATGATGACAA 4512
Db 4554 ATGAGCGGAGACTCAACTTAAGAGTCTGTAAACAGCTGAAGATGAAGATATGAG 4613
Qy 4513 GTAGACATTAATCTGTGACAGGTTCACCAACAGCTTTTGGCAGGACACAGGTTAC 4572
Db 4614 ATGACATAGATTAATGATGATTCATCAGACAGTGTCTTGAAGGAAAGGAGATT 4673
Qy 4573 TCGGTCAATGAGGCAATGTTATTTACCTGGAAGGTACAGATTCATCAGACGCGC 4632
Db 4674 AGTACTAAGAAAGAAATTTGATTTGATCTTGAAAGGACCAATTTCTATCAAGCAGCA 4733
Qy 4633 AAGGACATTCGCGAATTCATGACATGTGCGCAACAAATCTGAGCTAATGACAGATT 4692
Db 4734 AAGACATGCGGAGATTAAGTGTCTGTCTTAATGACAGGAAGAAATATGAACAATG 4793
Qy 4693 TGTCTTATCATCTCTGGGAGAGATATCCAGCATCCGCTTCAATTCAGATGAGAG 4752
Db 4794 TGTGCTTACATATTGGGTGAGACCATGGAAGCAATCCGCAAAAGTCCCGGTGACCAT 4853

```

OY 4753 TCAGAGGCGCTGCTGCTCACTTCCATGCTGTATTAATTAAGCTATGACGGCT 4812
    |||||
Db 4854 AACCCGTGCTTAGCCGCCCAAAAGTTGCCGTGCTTTGCATGATGCCATGACGCCA 4913
    |||||
OY 4813 GAGCGCGTATACAGAGTTGGCTCTGCGAAGAAAGAACAGTTGCCGTATGCTCATCATTC 4872
    |||||
Db 4914 GAAAGGGTCCACAGACTTAGAAGCAATACGTCAAGAAGTTACAGTATGCTCCTCCACC 4973
    |||||
OY 4873 CTGTGCCGAAGTACAGAGATCACAGGCGTGCAGAAAGCTACAGTGCAGCAAAACAGTCCCTG 4932
    |||||
Db 4974 CCCCTTCTTAAGCACAATAATTAGATGTTTCAGAAAGTTCAAGTGACGAAAGTAGTCCCTG 5033
    |||||
OY 4933 TTTTCAGGCGTGTACCAACCGCGTGTACACCCAGAGAAAGTACGCGAAAT 4982
    |||||
Db 5034 TTTAATCCGCACTCCGCGATTGCTCCCGCCGTAAGTACATAGAACT 5083
    |||||

```

Search completed: November 15, 2003, 20:02:56
 Job time : 382 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2003, 11:11:16 ; Search time 1177 Seconds
(without alignments)
11467.451 Million cell updates/sec

Title: US-10-023-649-1_COPY_1_5000
Perfect score: 5000
Sequence: 1 accctacaactaatcgtc.....ataattctagaacgcacc 5000

Scoring table:
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
- 25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4492	89.8	11492	20	AAV74107 Western equine enc
2	2118	42.4	11459	21	AAA49442 Venezuelan equine
3	2116.4	42.3	12523	24	AAD29131 Venezuelan equine
4	2113.2	42.3	12379	24	AAD29139 Human immunodefici
5	2113.2	42.3	13584	24	AAD29141 Venezuelan equine
6	2111	42.2	7479	24	AAD29132 Venezuelan equine
7	2074.8	41.5	11464	20	AAV74108 Venezuelan equine
8	1565.6	31.3	11060	20	AAV78130 pSVV link DNA. Se

9	1565.6	31.3	12464	21	AAC55473	Destination vector
10	1564	31.3	7399	25	ACC45139	Alphavirus replica
11	1564	31.3	15538	20	AAV76582	Plasmid pMP44 nucl
12	1562.4	31.2	11489	24	ABN86687	Nucleotide sequenc
13	1562.4	31.2	13599	24	ABN86690	Nucleotide sequenc
14	1554.4	31.1	11517	13	AAQ26021	pSP6-SFV4 RNA tran
15	1473.8	29.5	8010	20	AAV78129	Semliki Forest vir
16	1473.8	29.5	8100	20	AAV76588	SFV EcoRV-Spel tra
17	1333.2	26.7	11663	19	AAV33473	South African Arbo
18	1333.2	26.7	11663	19	AAV33475	South African Arbo
19	1331.6	26.6	11663	18	AAT47668	South African Arbo
20	1328.4	26.6	11717	21	AAA70577	Sindbis-like virus
21	1306.2	26.1	11626	21	AAV70607	Sindbis-like virus
22	1305.2	26.1	11703	21	AAC64506	Alphavirus SindCch
23	1305.2	26.1	11703	21	AAC64507	Alphavirus SindCch
24	1303.6	26.1	11703	19	AAV33476	Sindbis virus cion
25	1302.2	26.0	11717	19	AAV33474	Girdwood S.A.virus
26	1300.4	26.0	8000	25	ABX81525	Sindbis virus CDNA
27	1300.4	26.0	11740	25	ABX81526	Sindbis virus CDNA
28	1300.4	26.0	11927	20	AAV73356	Polynucleotide seq
29	1300.4	26.0	11927	21	AAV90388	Plasmid pSireps h
30	1298.8	26.0	9951	20	AAV77354	Polynucleotide seq
31	1298.8	26.0	9951	21	AAA90386	Sindbis expression
32	1298.8	26.0	9951	22	AAD04742	Alphaviral vector
33	1298.8	26.0	9951	24	ABV73174	Nucleotide sequenc
34	1298.8	26.0	9951	24	ABN86686	Nucleotide sequenc
35	1298.8	26.0	12110	24	ABN86689	Polynucleotide seq
36	1298.8	26.0	13905	20	AAV77358	Polynucleotide seq
37	1298.8	26.0	13905	21	AAA90390	Plasmid pRES 2J DN
38	1298.8	26.0	15958	25	AAI55269	DNA of expression
39	1297.2	25.9	10524	20	AAV73355	Polynucleotide seq
40	1297.2	25.9	10524	21	AAA90387	Plasmid pSireps E
41	1295.6	25.9	8000	18	AAV58321	Sindbis virus vari
42	1295.6	25.9	8000	20	AAV58571	Sindbis virus vari
43	1295.6	25.9	8000	25	ABX81524	Sindbis virus CDNA
44	1295.6	25.9	11282	20	AAZ27850	Complete CDNA sequ
45	1295.6	25.9	11282	22	AAF84024	Complete CDNA sequ

ALIGNMENTS

RESULT 1	
AAV74107	standard: CDNA, 11492 BP.
AC	AAV74107;
AC	AAV74107;
DT	12-APR-1999 (first entry)
DE	Western equine encephalitis virus CDNA.
XX	
KW	WEE virus; vaccine; ds.
OS	Western equine encephalitis virus.
XX	
FN	MO9853077-A1.
PD	26-NOV-1998.
XX	
PF	20-MAY-1998; 98MO-US10645.
XX	
PR	16-DEC-1997; 97US-0991840.
PR	20-MAY-1997; 97US-0047162.
PR	24-JUL-1997; 97US-0053652.
XX	
PA	(REED-) REED ARMY INST RES WALTER.
PA	
PI	Crisse BJ, Oberste MS, Parker MD, Schmura SM, Smith JF;
XX	
DR	WPI, 1999-045316/04.
XX	
PT	New DNA encoding infectious Western or Venezuelan equine

PT encephalitis virus genome - useful for the production of live or
PT attenuated vaccines for human or veterinary medicine

PS Example 1; Page 54-67; 112pp; English.

XX This cDNA sequence codes for an infectious western equine
CC encephalitis (WEE) virus RNA transcript. DNA representing the
CC entire genome was prepared by PCR using primers (see AAV74110-21)
CC based on partial genome sequences. 5' Sequences were obtained by
CC RACE. The full-length infectious clone is useful in the production
CC of virulent WEE virus, and for introducing and testing attenuating
CC mutations. Also new are: infectious or attenuated WEE RNA
CC transcripts and WEE viral particles; cDNA (see AAV74108) encoding an
CC infectious Venezuelan equine encephalitis (VEE) variant IE viral
CC genome; infectious or attenuated VEE RNA transcripts and viral
CC particles; an attenuated chimeric virus containing non-structural
CC sequences from a first alpha-virus (AV) and structural sequences
CC from a second AV, resulting in attenuation of the second AV; a method
CC for expressing a protein (especially an antigen to protect against
CC a pathogen) by cloning its gene into an attenuated WEE or VEE IE
CC replicon so that transcription of the replicon produces RNA able to
CC infect the cells in which protein is to be produced; methods for
CC diagnosing WEE and VEE infections; and polypeptides encoded by VEE
CC variant IIIA. Attenuated WEE and VEE are used in live or inactivated
CC vaccines, for use in human or veterinary medicine. Chimeric viruses
CC are also useful as vaccines, directed against the second AV which is
CC particularly Eastern equine encephalitis (EEE) virus or some variant
CC of VEE. WEE and VEE nucleic acids are used as primers and probes to
CC diagnose virus infections and to define natural variants, also for
CC production of protein antigens which can be used as diagnostic
CC reagents, to generate antibodies, and in vaccines. The attenuated
CC viruses are highly immunogenic and provide long-lasting protection.

XX Sequence 11492 BP; 3257 A; 2854 C; 2794 G; 2570 T; 17 other:

Query Match 89.8%; Score 4492; DB 20; Length 11492;

Best Local Similarity 95.8%; Pred. No. 0;

Matches 4826; Conservative 0; Mismatches 127; Indels 82; Gaps 18;

QY 1 ACCCTCAAACTAATTCGATCCATATGGAAGAATTCCGTTGACTTATGCTGACAC 60
DB 25 ACCCTCAAACTAATTCGATCCATATGGAAGAATTCCGTTGACTTATGCTGACAC 84
QY 61 CCGATGTCAGTGTACAGGAGCGTTTCCACATTTGATGCAAGCAAGGAGGTC 120
DB 85 CCAATATGTCAGTGTACAGGAGCGTTTCCACATTTGATGCAAGCAAGGAGGTC 144
QY 121 ACTGACATGACATGCGCAATGCGAGCGTTTTCGATGTCGCAACAAAGCTATTGAG 180
DB 145 ACTGACATGACATGCGCAATGCGAGCGTTTTCGATGTCGCAACAAAGCTATTGAG 204
QY 181 AGCGAATGTCAGCGGAGCAAGTTATCTTGACATTTGGAAGTGGCCCGCTGACATGCA 240
DB 205 AGCGAATGTCAGCGGAGCAAGTTATCTTGACATTTGGAAGTGGCCCGCTGACATGCA 264
QY 241 CATTCACATGACCGCTATCATTTGATCTGCGCTATGATTAAGCGTGAAGACCCGAGACA 300
DB 265 CATTCACATGACCGCTATCATTTGATCTGCGCGCTATGATTAAGCGTGAAGACCCGAGACA 324
QY 301 CTACACGGTATGAGAAAGACTTAAGAAAGTGAACAT-TACGACAAAGACATAGCTC 359
DB 325 CTACACGGTATGAGAAAGACTTAAGAAAGTGAACAT-TACGACAAAGACATAGCTC 384
QY 360 T-AAGGGGGGAGACCTGCTGGAAGTCAATGCAACGACGACGAGACTTCATCTCTGT 418
DB 385 TNAAGGGGGGAGACCTGCTGGAAGTCAATGCAACGACGACGAGACTTCATCTCTGT 444
QY 419 GTATGACACAGAGCAAGTGTAGTACTTTGGAAGTGA-TGAGTATTAACAAGATGTG 477
DB 445 GTATGACACAGAGCAAGTGTAGTACTTTGGAAGTGA-TGAGTATTAACAAGATGTG 504
QY 478 TACGACATGTCATGACGACATCAATCTACACGAGGCGCTTAAGAGATTAGACATTT 537

DB 505 TAGCAGATCCATGACCCGACATCAATCTACACGAGCGCTAAAGAGTTAGACATTT 564
QY 538 TACTGATAGGCTTTGACAGACCCCTTTATGTAACAAAACATGGAGGTTCTTACCCT 597
DB 565 TACTGATAGGCTTTGACAGACCCCTTTATGTAACAAAACATGGAGGTTCTTACCCT 624
QY 598 ACTTACACAGCAAGCTGGGCTGACAGAGAGTATTGGAACGATACATTTGGCTCGGT 657
DB 625 ACTTACACACAACTGGGCTGACAGAGAGTATTGGAACGATACATTTGGCTCGGT 684
QY 658 AACTCAGATCTTACAGAGAGAGCGTTGAAAACCTCAATCCTTACGAAAGAGGCTC 717
DB 685 AACTCAGATCTTACAGAGAGAGCGTTGAAAACCTTCAATCCTTACGAAAGAGGCTC 744
QY 718 CAACCTTCTAATAAGATCAATTTCTGGTGTTCACATCTACAGAAAGATAGTCA 777
DB 745 CAACCTTCTAATAAGATCAATTTCTGGTGTTCACATCTACAGAAAGATAGTCA 804
QY 778 CTGTTACGTAGCTGGCATCTTCCAAACGTTTCCACTTGAAGAAAGTCACTTACACA 837
DB 805 CTGTTACGTAGCTGGCATCTTCCAAACGTTTCCACTTGAAGAAAGTCACTTACACA 864
QY 838 GTATGATGTGGACCATTTGTACGCTGTAAGGGTACGTCATCAAAAAGATPACGATCAGC 897
DB 865 GTATGATGTGGACCATTTGTACGCTGTAAGGGTATGTCATCAAAAAGATPACGATCAGC 924
QY 898 CCAGGATTAATACGTTAAAGTTGAAACTTGGCGTCCCAATGATCCGAGGGTTCTTG 957
DB 925 CCAGGATTAATACGTTAAAGTTGAAACTTGGCGTCCCAATGATCCGAGGGTTCTTG 984
QY 958 AGTTGCAAAATCAAGATACGTCGCGGCGAGAGGGTTTCTTTGCTGTGTACGTAT 1017
DB 985 AGTTGCAAAATCAAGATACGTCGCGGCGAGAGGGTTTCTTTGCTGTGTGTACGTAT 1044
QY 1018 GTACACGCCACACTTTCGATCAGATGACAGGATTCGTGCAACTGATGTGAGT 1077
DB 1045 GTACACGCCACACTTTCGATCAGATGACAGGATTCGTGCAACTGATGTGAGT 1078
QY 1078 GAGCGCAAAAACATTTGTTGGCTTCAACAAAGATTTGTCATATGTTAGAGCGCAA 1137
DB 1079 -----GGGCTCAACCAAGATTTGTCATATGTTAGAGCGCAA 1117
QY 1138 AGAATATCTAACAACAATGACAACTATCTATTACAGTGTGCGCCAGCGTTTCCAGG 1197
DB 1118 AGAATATCTAACAACAATGACAACTATCTATTACAGTGTGCGCCAGCGTTTCCAGG 1177
QY 1198 TGGCGCGTGAACATCGTGCAGCTTGAGCAGCAGAAAGAACTAGGGGTGCGGAGCGC 1257
DB 1178 TGGCGCGTGAACATCGTGCAGCTTGAGCAGCAGAAAGAACTAGGGGTGCGGAGCGC 1237
QY 1258 ACTCTTAATATGCGCTGCTGCTGGCTTCAAGCCGAAATTCATCATCTTACAG 1317
DB 1238 ACTCTTAATATGCGCTGCTGCTGGCTTCAAGCCGAAATTCATCATCTTACAG 1297
QY 1318 AAGCGTGTGAGCAAAAC-AAATTAAGAAAGTACCTGCGCTC-TTGTACTCTTTGTGATTC 1375
DB 1298 AAGCGTGTGAGCAAAAC-AAATTAAGAAAGTACCTGCGCTC-TTGTACTCTTTGTGATTC 1357
QY 1376 C-ACGCTTACACAGCCAGCGGCTCGAT-----ATGGGCTTCCGCGGT-AGGCTCAAGCT 1427
DB 1358 C-ACGCTTACACAGCCAGCGGCTCGAT-----ATGGGCTTCCGCGGT-AGGCTCAAGCT 1417
QY 1428 GCTGCTTGAACCAACTGTCGAACCCGACCGGCTATTAACAATGCGGATGAGGAGCATCT 1487
DB 1418 GCTGCTTGAACCAACTGTCGAACCCGACCGGCTATTAACAATGCGGATGAGGAGCATCT 1477
QY 1488 GCGTGGCTTACAGAAAGACTGAAAGAGTGTGTCAGC-GGAAGAGATCAGAGAAAGCCC 1546
DB 1478 GCGTGGCTTACAGAAAGACTGAAAGAGTGTGTCAGC-GGAAGAGATCAGAGAAAGCCC 1537
QY 1547 TGGCACCCCTTGTCTCCCTGAATATGAAGAAAGAGACCGTGAAGGAGAAAGTACCTCATTA 1606
DB 1538 TGGCACCCCTTGTCTCCCTGAATATGAAGAAAGAGACCGTGAAGGAGAAAGTACCTCATTA 1597

QY 1607 TCACAGAGCAGAGCAGGTAGCGTGAGACAC-ACGAGACACATCAGGCTGACAAGT 1665
DB 1598 TGCAGAGGCGAGAGAGGTAGCGTGAGACACACACGAGACATATCAGGTTACAAAGT 1657
QY 1666 TACCCAGGCG-AAGAGAGATGGGTCTTA-CGCTATCTTTCACCCAGGCGGATTT-G 1722
DB 1658 TACCCAGGCGAGAGAGATGGGTCTTATCCCTATCTTTCACCCAGGCGGATTTTAN 1717
QY 1723 AATAGTAAAACTGGCGGTATCCACCATTTGGGGAAACAAGTCTGGTATGACTCAC 1782
DB 1718 AATAGTAAAACTGGCGGTATCCACCATTTGGGGAAACAAGTCTGGTATGACTCAC 1777
QY 1783 AAAGTAGGGCAGGAGAGATACAAAGTCAGCCATACACGCTAAGTCACTTTGA-CCAGA 1841
DB 1778 AAAGCAGGGCGGGGGATACAAAGTCAGCCATACACGCTAAGTCACTTTGAACCCAGA 1837
QY 1842 AGGACGGC-GGTCCCTGTTCAAGACTTCAGGCACTTGAAGTAGAGCGCTACGATCGTT 1900
DB 1838 AGGACGGCGGGTCCCTGTTCAAGACTTTCAGGCACTGAAGTAGAGCGCCAGATCGTTT 1897
QY 1901 TCAAGAGAGGAGTTCGTA------CAGATACCTGCACCCATCCGATTA 1948
DB 1898 TCAAGAGAGGAGTTCGTA-CCAGATATTTGACCCACATCCGACCKTTTCAACTA 1957
QY 1949 ACGAGAGCGGCTAAACACTGACGAAAGTACTATAGACTGTAAAGACTCAGAGACAG 2008
DB 1958 TAGTAGTCGCTATTAAGTACGAGAGTACTATAGTGTAAAGACTCAGAGAGCGAG 2017
QY 2009 ACTCAGAAATCGTCTTGCATATTAAGCGACGAAAGTGTGTTAAGCGAGAAAGCGAGGTC 2068
DB 2018 ACTCAGAAATCGTCTTGCATTAAGCGACGAAAGTGTGTTAAGCGAGAAAGCGAGGTC 2077
QY 2069 CCTTGCTTAAACCGGTGATCTGTAATCCACATTTACAGAGTTTGGTACGAGAGTC 2128
DB 2078 CCTTGCTTAAACCGGTGATCTGTAATCCACATTTACAGAGTTTGGTACGAGAGTC 2137
QY 2129 TCAAGACAGCAGCAGCAGCCTCAAAAGTCCCAACCATCGAGTCTATGAGAGCGAG 2188
DB 2138 TCAAGACAGCAGCAGCAGCCTCAAAAGTCCCAACCATCGAGTCTATGAGAGTCAG 2197
QY 2189 GTTCAGGTAAATCTGGAATCATCAAAAGCGCTGTGACTAAGAAAGTCTGTTGATG 2248
DB 2198 GTTCAGGTAAATCTGGAATCATCAAAAGCGCTGTGACTAAGAAAGTCTGTTGATG 2257
QY 2249 CGAAGAGGAAATCTGCGAGAAATCATAGGAGTGTAAAGAGTGAAGCTATGATG 2308
DB 2258 CGAAGAGGAAATCTGCGAGAAATCATAGGAGTGTAAAGAGTGAAGCTATGATG 2317
QY 2309 TTGCTGCTAGAGCTGCTGATCAGTCTCTAAATGGGGTTAAGCACCCTTTACACTC 2368
DB 2318 TTGCTGCTAGAGCTGCTGATCAGTCTCTAAATGGGGTTAAGCACCCTTTACACTC 2377
QY 2369 TGTACATTTGATGAGCATTTGCTGCTGATGAGGAGCGCTGCTGATGATGTCATG 2428
DB 2378 TGTACATTTGATGAGCATTTGCTGCTGATGAGGAGCGCTGCTGATGATGTCATG 2437
QY 2429 TCAAACTTAAGAAAGTGTATGTTGGGGGAGCCCAAAACATGCGGCTTTTAACATGA 2488
DB 2438 TCAAACTTAAGAAAGTGTATGTTGGGGGAGCCCAAAACATGCGGCTTTTAACATGA 2497
QY 2489 TGTGCTGAAAGTACATTTTAACATGACATGACATGACATGACATGACATGACATG 2548
DB 2498 TGTGCTGAAAGTACATTTTAACATGACATGACATGACATGACATGACATGACATG 2557
QY 2549 CTAGAGAGTGCACACAGACTGTAAACCGCATGCTTCAACGCTCTTCAACGACGAA 2608
DB 2558 CTAGAGAGTGCACACAGACTGTAAACCGCATGCTTCAACGCTCTTCAACGACGAA 2617
QY 2609 TGAAGACGTTAAACCATGCTGATTAATCATATGATATCCACAGGAGACCAAAAGC 2668
DB 2618 TGAAGACGTTAAACCATGCTGATTAATCATATGATATCCACAGGAGACCAAAAGC 2677

QY 2669 CGCACAAAGATGATCTGATTTCAACCTGTTTCAGAGATGGGTGAAGACGTACAGATTG 2728
DB 2678 CGCACAAAGATGATGATTTCAACCTGTTTCAGAGATGGGTGAAGACGTACAGATTG 2737
QY 2729 ACTACAAAATCAAGAAATCATGACTGCGGCTGCATCCGAAAGACTTACGCGGAAAGCG 2788
DB 2738 ACTACAAAATCATGAAATCATGACTGCGGCTGCATCCGAAAGACTTACGCGGAAAGCG 2797
QY 2789 TTTATGCTGACAGTACAAAGTCAAGGAATCCACTGATCTGCGAGCTTCTGAGCACG 2848
DB 2798 TTTATGCTGACAGTACAAAGTCAAGGAATCCACTGATCTGCGAGCTTCTGAGCACG 2857
QY 2849 TGAAGCTTACTTACACGACAGAAAAACGATTTGTGGAAGAAGCTAGCTGTGATC 2908
DB 2858 TGAAGCTTACTTACACGACAGAAAAACGATTTGTGGAAGAAGCTAGCTGTGATC 2917
QY 2909 CCTGATTAAGACACTTACAGTAAATATCCCGGGGATTTCACGGCTTCAITGACGACT 2968
DB 2918 CCTGATTAAGACACTTACAGTAAATATCCCGGGGATTTCACGGCTTCAITGACGACT 2977
QY 2969 GGGACGGGAAACAGCGCCATTTAAGACGGGTTCTGATTAAGCGGAGACAGCTGATG 3028
DB 2978 GGGACGGGAAACAGCGCCATTTAAGACGGGTTCTGATTAAGCGGAGACAGCTGATG 3037
QY 3029 TGTTCAGAAATTAAGGTGAACGTCCTGCGGAGAGGCTTTAAGAGCCAGTCTTGCGCACG 3088
DB 3038 TGTTCAGAAATTAAGGTGAACGTCCTGCGGAGAGGCTTTAAGAGCCAGTCTTGCGCACG 3097
QY 3089 CCAACATTTGCTGACGAGACAGCAGTGGAGAGCTTGCCACCATTTCAAGCATGACAG 3148
DB 3098 CCAACATTTGCTGACGAGACAGCAGTGGAGAGCTTGCCACCATTTCAAGCATGACAG 3157
QY 3149 CGTACTCACTGAATGAGCACTGAACCTTTTGAACAGAGTCTTGAAGTACAGACCTG 3208
DB 3158 CGTACTCACTGAATGAGCACTGAACCTTTTGAACAGAGTCTTGAAGTACAGACCTG 3217
QY 3209 ACAGTGGGTATTTTCCGCTCCTACCGTCGACACTTACTTACAGGAGTACGACTGGATA 3268
DB 3218 ACAGTGGGTATTTTCCGCTCCTACCGTCGACACTTACTTACAGGAGTACGACTGGATA 3277
QY 3269 ACTGCCAGGAGAAATGTAAGGCTTTAATGAGAGTACAAAGAGTGTGACAGG 3328
DB 3278 ACTGCCAGGAGAAATGTAAGGCTTTAATGAGAGTACAAAGAGTGTGACAGG 3337
QY 3329 GATATCCGTGATCACAAAACCGGTTGACAGAGAGGTAGCTGATTAAGAAATTA 3388
DB 3338 GATATCCGTGATCACAAAACCGGTTGACAGAGAGGTAGCTGATTAAGAAATTA 3397
QY 3389 CCATCAAGAGTACTCTCCCAAAATTAATGTTGTTCCATTAAATGCGCGTTGCCCACT 3448
DB 3398 CCATCAAGAGTACTCTCCCAAAATTAATGTTGTTCCATTAAATGCGCGTTGCCCACT 3457
QY 3449 CGTTGATGTTGACCAAAAGAGCAGGGTACACTGATCAGAGCGGATTCCTATCTAAGA 3508
DB 3458 CGTTGATGTTGACCAAAAGAGCAGGGTACACTGATCAGAGCGGATTCCTATCTAAGA 3517
QY 3509 TGAAGGCAAAATCTGTGTTGATGCGGATCCATACAGACTTCAGGAGAAAGTAG 3568
DB 3518 TGAAGGCAAAATCTGTGTTGATGCGGATCCATACAGACTTCCTGGGAAAGAAAGTAG 3577
QY 3569 AGTCCATGGGTCAATGCGCACTAATACATCAGAGTGTGATCGATTTGGGAATACCTA 3628
DB 3578 AGTCCATGGGTCAATGCGCACTAATACATCAGAGTGTGATCGATTTGGGAATACCTA 3637
QY 3629 GCCATGCGGTAAATATGACATTAATCTTGTCAATGTTAGACCCCGTACAGAAACATC 3688
DB 3638 GCCATGCGGTAAATATGACATTAATCTTGTCAATGTTAGACCCCGTACAGAAACATC 3697
QY 3689 ACTACCAACAGTGGAGAGATACGCTATCCACACAGATGCTAAACGTGTAAGGCTGTCC 3748
DB 3698 ACTACCAACAGTGGAGAGATACGCTATCCACACAGATGCTAAACGTGTAAGGCTGTCC 3757
QY 3749 ACCACCTGAACATGTCGGAGAACATGTCGATATGAGGTATGAGGCTCTGATCGCGCA 3808

```
Db 3758 ACCACCTGGAACACTGGGGAAACATGTGTGSCCATAGGTTATGACCTTGTCTATGCGCGAA 3817
Qy 3809 CCGAGAAATATCATCATCTGCGGTGCGACGCTCATTTAGTTTACCCTGTCTGTACGCTTA 3868
Db 3818 CCGGAAATATCATCATCTGCGGTGCGCTGCTCATTTAGTTTACCCTGTCTGTACGCTTA 3877
Qy 3869 AGAACATGCGCGGAAATATCTAGAGTCTCTGTTCTTGGGCAAGGACAAACGCGAAC 3928
Db 3878 AGAACATGCGCGGAAATATCTAGAGTCTCTGTTCTTGGGCAAGGACAAACGCGAAC 3937
Qy 3929 ACACACATGACACAGACAGACTCGGTGTAGTGTCTTGAACAATCATCAAGGTCACCA 3988
Db 3938 ACACACATGACACAGACAGACTCGGTGTAGTGTCTTGAACAATCATCAAGGTCACCA 3997
Qy 3989 GGTACGAGGACAGGAGAGCTCCAGGCTACAGAGTATCAGAGTGCATTTAGCAAGGCG 4048
Db 3998 GGTACGAGGACAGGAGAGCTCCAGGCTACAGAGTATCAGAGTGCATTTAGCAAGGCG 4057
Qy 4049 CTGACCAAGCTATGTTAATGCTGTATATAGCAAAAGGTCACAGGTTCCGGAGTGTGCG 4108
Db 4058 CTGACCAAGCTATGTTAATGCTGTATATAGCAAAAGGTCACAGGTTCCGGAGTGTGCG 4117
Qy 4109 GTGCACTGTACCGAAATGCGCGCTGCTTTGTATAGACAGCCAAATAGCTGTCCGGAACG 4168
Db 4118 GTGCACTGTACCGAAATGCGCGCTGCTTTGTATAGACAGCCAAATAGCTGTCCGGAACG 4177
Qy 4169 CTAGACTTTGAGACGACGACCGCTCATCATCATCTGTGTAGAACCCAAATTTTCTAAGA 4228
Db 4178 CTAGACTTTGAGACGACGACCGCTCATCATCATCTGTGTAGAACCCAAATTTTCTAAGA 4237
Qy 4229 TGCCGGAAACCGGAGGCGGACCTTAAGCTGCGAGCGCTCATATAGCAATACGTCACATCG 4288
Db 4238 TGCCGGAAACCGGAGGCGGACCTTAAGCTGCGAGCGCTCATATAGCAATACGTCACATCG 4297
Qy 4289 TCACACGCTGAGCGGATTCACAAAATATCATAGTACCGCTACTGTCAACCGGCACTTATCTG 4348
Db 4298 TCACACGCTGAGCGGATTCACAAAATATCATAGTACCGCTACTGTCAACCGGCACTTATCTG 4357
Qy 4349 GTGCGAAAGATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4408
Db 4358 GTGCGAAAGATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4417
Qy 4409 ATGCGAGTGTACCATATATGCTTGTGATTAACAATGAGGACACAGATTAATGAGGCGCA 4468
Db 4418 ATGCGAGTGTACCATATATGCTTGTGATTAACAATGAGGACACAGATTAATGAGGCGCA 4477
Qy 4469 TTACACCGCAAGAAAGCTGCGAAATTTCTG--GATGATGACAAAGCAGTATGACATTTGACT 4525
Db 4478 TTACACCGCAAGAAAGCTGCGAAATTTCTG--GATGATGACAAAGCAGTATGACATTTGACT 4537
Qy 4526 TGTGACGGGTCCACCCAAACAGCTTTTGGGACGACGACGAGTTTCTCCGTCAATGAGG 4585
Db 4538 TGTGACGGGTCCACCCAAACAGCTTTTGGGACGACGACGAGTTTCTCCGTCAATGAGG 4597
Qy 4586 GCAAGTTTATTCATCTGAGAGGTACAGATTCATCAACGCGCAAGGACAACTTGGCG 4645
Db 4598 GCAAGTTTATTCATCTGAGAGGTACAGATTCATCAACGCGCAAGGACAACTTGGCG 4657
Qy 4646 AAATTCATGCAATGTGCGCAACAAATCTGAGGCTTAATGACAGATTTGCTTGTATCATCC 4705
Db 4658 AAATTCATGCAATGTGCGCAACAAATCTGAGGCTTAATGACAGATTTGCTTGTATCATCC 4717
Qy 4706 TGGGGAGAGTATGTCCAGCATCCGCTCCAAATGCGCGTAGAGAGTACAGAGGCGTCTG 4765
Db 4718 TGGGGAGAGTATGTCCAGCATCCGCTCCAAATGCGCGTAGAGAGTACAGAGGCGTCTG 4777
Qy 4766 CTACACCTCACAACCTTCATGCGCGGTGTATTAATGATGACGCGCTGAGCGCGATTAACA 4825
Db 4778 CTACACCTCACAACCTTCATGCGCGGTGTATTAATGATGACGCGCTGAGCGCGATTAACA 4837
Qy 4826 GGTTCGCTGTGCGAAGAAAGACAGTTGCGCGTATGCTCATCTCTGTTGCGGAAGT 4885
```

```
Db 4838 GGTTCGCTGTGCGAAGAAAGACAGTTGCGCGGTATGCTCATCTCCGTCGCGAAGT 4897
Qy 4886 ACAGATACACAGCGGTGACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGT 4945
Db 4898 ACAGATACACAGCGGTGACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGT 4957
Qy 4946 TACACCGGCTGTACACCCAGGAGTACCGGAAATATTTAGAAACCCAC 5000
Db 4958 TACACCGGCTGTACACCCAGGAGTACCGGAAATATTTAGAAACCCAC 5012

RESULT 2
AAA49442
ID AAA49442 standard; DNA; 11459 BP.
XX
AC AAA49442;
XX
DT 25-SEP-2000 (first entry)
XX
DE Venezuelan equine encephalitis virus (VEE) wild type sequence.
XX
KM Venezuelan equine encephalitis virus; VEE; attenuation; vaccine;
KM replicon; heterologous gene expression; capsid; structural protein;
KM glycoprotein; immunogen; ds.
XX
OS Venezuelan equine encephalitis.
XX
FN MO200034445-A2.
XX
PD 15-JUN-2000.
XX
PE 07-DEC-1999; 99WO-US29041.
XX
PR 07-DEC-1998; 98US-0111330.
XX
PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
PI Pushko P, Parker MD, Smith JF, Crise BJ;
XX
WP1: 2000-423405/36.
XX
PT Attenuated Venezuelan equine encephalitis virus useful as a vaccine
PT contains a rearrangement in genome such that its structural
PT glycoproteins precede the capsid gene
XX
PS Disclosure; Page 37-43; 43pp; English.
XX
XX
CC An attenuated Venezuelan equine encephalitis (VEE) has been described
CC in which the virus undergoes a rearrangement of its genome such that
CC the structural glycoproteins precede the capsid gene. The resulting
CC attenuated virus is useful as a live VEE vaccine for human and
CC veterinary use. The attenuated virus can also be used in a replicon
CC expression system. The rearranged VEE replicon can be used to express
CC heterologous genes of interest in cells as well as a means for
CC expressing antigens or immunogenic proteins and peptides of interest.
CC This rearranged sequence appears in the attenuated virus designated
CC RV1.1.1.
XX
SQ Sequence 11459 BP; 3242 A; 2804 C; 2897 G; 2516 T; 0 other;
XX

Query Match 42.4%; Score 2118; DB 21; Length 11459;
Best Local Similarity 64.5%; Pred. No. 0;
Matches 3219; Conservative 0; Mismatches 1745; Indels 24; Gaps 3;
```

Db 150 GACCAATGCTAAAGCCAGAGCGTTTTCGCACTCGCTTCAAACTGATCAACGAGG 209
Qy 190 GACCGGACCAAGTTATCTTGACATTTGGAGTGGCCCGGTCAGACATGCAATTCAT 249
Db 210 GACCATCCGACGACATCTTGAACATTTGAGTGGCCCGGTCAGAAATGTAATCTTA 269
Qy 250 CACCCATATCTATCTGCTCCATGATTAAGCGCTGAAGACCCGGACAGATCAACAG 309
Db 270 CACAGTATCATATGATCTGCTCCATGATGATGGGAGATCCGACAGATTTGATAG 329
Qy 310 TATGCAAGAAAGCTTAAAGAAAGT-----GACATTAACGACAGAAACATAGCTTAA 363
Db 330 TATGCAACTAAGCTAAGAAAGAAAGTAAAGAAATTAAGTAAAGAAATGAGCAAGAA 389
Qy 364 GCGGAGACCTGCTGGAAGTCAATGTCACACGAGACGAGATCCATCTCTGTGTATG 423
Db 390 ATGAAGAGAGCTGCGCGCTCATGAGCGACCTGACCTGGAACCTGAGACTATGTGCTC 449
Qy 424 CACAGACGCGACGCTGATGACTTTGGAGTGTAGCAGTATACCAAGATGTATACGA 483
Db 450 CACGACGAGAGAGTGTGTGCTACGAAAGGCAAGTGTGCTTTACAGAGATTTATACCG 509
Qy 484 GTCCATGCAACGACATCAATCTTACCAACGCGCTTAAAGAGTTAGACAAATTTACTG 543
Db 510 GTTGACGACCGACAGTCTCTATCACCAAGCCAAATTAAGGAGTTAGAGTGCCTACTG 569
Qy 544 ATAGCTTTGACACGACCCCTTTTATGTAACAAAACATGGCAGTTCCTACCTACTAC 603
Db 570 ATAGCTTTGACACGACCCCTTTTATGTAACAAAACATGGCAGTTCCTACCTACTAC 629
Qy 604 AACAGAACTGGGCTGACGAGAGATTTGAAAGACGTAATTCATGGCTCGGTAACTCA 663
Db 630 TCTACCACTGGGCGGACGAAACCGTGTAAAGGCTCTGTAATAGGCTTATGCACTCT 689
Qy 664 GATCTTCAGAGAGAGCGCTTGAAGAACTCTCAATCTTTAGAGAGAGGCTCCAACT 723
Db 690 GACGTTATGAGAGCGGTCAAGTAGAGAGATGTCATTTCTTAAGAAAGATTTGAAACA 749
Qy 724 ACTATATAGATCATATTTCTCGGTTGTTCAACATCTTACAGAAATGATGACTGTTA 783
Db 750 TCCAACTATGTTCTATTTCTGTGTGCTGACCATCTTCAACAGAAAGGAGCTTACTG 809
Qy 784 CGTACCTGGCATCTTCCAAACGTTTCCACTTGAAGGAAAGTTAACTTCAAGAGTGA 843
Db 810 AGGAGCTGGACCTCGCTCTATTTCACTTACGTCGACAAATTTACACATGTGCG 869
Qy 844 TGTGGACCATTTGTCAGCTGTGAGAGGTAAGTATCAAAAAGATTAACATCAGCCAGA 903
Db 870 TGTGAGACTATAGTTAGTTGCGAGGTAAGTGTAAAGAAATGCTATCAGTCAAGC 929
Qy 904 CTATACGCTAAAGTTGAGAACTTGGCGTTCACAAATGCAATCGGAGGCTTTCTTGA 963
Db 930 CTGATGAGGAGCCCTTCAAGGCTATGCTGTACGATGCAACCGGAGAGGATTTCTGTG 989
Qy 964 AAAGTCACAGATAGCGCTGCGGCGAGAGGCTTTCTTTGCTGTGATGATGATACCA 1023
Db 990 AAAGTCACAGACATTTGAACGGGAGAGGCTTCTTTTCCGCTGTGACGATGTGCA 1049
Qy 1024 GCCACACTTTGGCATCAGATGACAGGATTTGGCAACTGACGTTAGTGTGATGACCA 1083
Db 1050 GCTACATTTGTGACCAATATGACTGCAATGCGCAACGATGTCAGTGGGAGACGCGG 1109
Qy 1084 CAAAAAATATTTGTTGGGCTCAACCAAGAGATTTGCTCAATGTAAGACGCAAGAAAT 1143
Db 1110 CAAAAAATGCTGTGTGGCTCAACCAAGGATATGCTCAACGCTGCGACCCAGAGAAAC 1169
Qy 1144 ACTACACAATGACAGAACTATTTACCAAGTGTGCGCCAGGCTTTTCCAGTGGGCG 1203
Db 1170 ACCAATACATGAAAAATTAATCTTTTCCGTAAGTGGCCAGGCAATTTGCTAGTGGG 1229
Qy 1204 CGTGAACATCGTGCAGCTTGGACGACGAGAAAGAACTAGGGGTGCGGAGCGCACTCT 1263

Db 1230 AAGGAATATAAGGAAGATCAAGAGATGAAGGCCACTAGAGCTACGATAGACATTA 1289
Qy 1264 ACTATGGCTGCTGTGGGCTTTTCAAGACCCAGAAATACATCCATGACAAGAGCT 1323
Db 1290 GTCATGGGCTGTTGTTGGCTTTTGAAGGACAAAGATTAACATTTATTAAGGCCCG 1349
Qy 1324 GTATGCAAAATTTAAGAAAGTACCTGCTCTTTGACTGATTTGTATTCACGCTT 1383
Db 1350 GATACCAAAACATCATTAAGTGAAGAGCATTTCCACTATTTGTCTGCTGCCAGATTA 1409
Qy 1384 ACCAGCCAGGAGCTGATATGAGCTTCCGCGTGAAGCTCAAGCTGCTCTGAACCACT 1443
Db 1410 GGCAGTAAACATTTGAGATGAGGCTGAGAACAGAAATCAGGAAATGTTAGAGAGAC 1469
Qy 1444 GTCAAAACCCGACCGGCTATTACAAATGCGCATGAGAGCATCTGCGGTGCTTACAGCA 1503
Db 1470 AAGAGCGCTACCTCTCATTTACCGCCGAGACGTAACAAGAGCTTAAGTGGCCAGT 1529
Qy 1504 GAAGCTGAAGAAAGTGGCTGACGCGAGAGATCAGAGAAAGCTTGCACCTTGTCTCCT 1563
Db 1530 GAGGCTTAAGAGAGTGTGCGTAAGCGAGAGATTTGGCGCAGCTTACCACTTTGGCAGT 1589
Qy 1564 GAATATGAAGAAAGAACCTGTAGAGCAGAGTGAACCTCATTTATGCAAGAGCAGACA 1623
Db 1590 GATGTTGAGAGCCCACTGTGAAAGCCGATGTGCACTTGAATGTTACAGAGGCTGGGCG 1649
Qy 1624 GGTAGCGTGAACCAACGAGAGACATCTAGGGTGAACAAGTTACCCGAGCGGAGAGAG 1683
Db 1650 GACTCAGTGAAGACCTCTGTGCTTGTAAAGTTACACTAGCTGCGGAGAGACAG 1709
Qy 1684 ATGGGCTTACGCTATCTTTCAACCCGAGCGGATTTGAATGTAAGAAATCTGGCTGT 1743
Db 1710 ATCGCTTTACGCTGTCTTTCGCGAGGCTGTACTCAAGAGTAAATTAATCTTGC 1769
Qy 1744 ATCCACCATTTGGGGAACAGTACTGTAAATGACTCAAAAGTAAAGGAGAGATAC 1803
Db 1770 ATCCACCTCTGCTGAAACAGTCAATGTAATACACTCTGCGCGAAAGGCGTTAT 1829
Qy 1804 AAAGTCGACCATCAAGGTAAGTCAATTTACAGAGAGGAGCGGCTCCGTGTA 1863
Db 1830 GCCGTAAGCAATCAAGTGAAGTGTGTGCGAGAGGACATGCAATCCGCTCAG 1889
Qy 1864 GACTTCCAGGATGAGTGAAGCGCTTCAATGCTTTCAACGAGAGGAGTCTGTAAC 1923
Db 1890 GACTTCAAGCTCTGAGTGAAGTGTCCACATTTGTGTCAACGAACTGATGCTTAAC 1949
Qy 1924 AGATACCTGCAACATGCAATCAACGAGAGCGCTAAACACTGACGAAGATATAT 1983
Db 1950 AGTATCTGCAACATATTTGCCACACATGAGAGGCTGAACACTGATGAAGATATAT 2009
Qy 1984 AAGACTGTAAGACTCAGAGACACAGACTCAGAAATACGTTCTGTGATTTGACGACGAAG 2043
Db 2010 AAAACTGTCAAGCCGAGGACACAGACGCGCAATACCTGTACGATGACAGGAAACG 2069
Qy 2044 TGTGTTAAGCAGAAAGACGAGTCCCTTGTGCTTAAACGAGATCTGTGATGATACCA 2103
Db 2070 TCGCTCAAGAAAGATCTGTACTGTGGCTTACAGGCGAGCTGTGATCTCTCC 2129
Qy 2104 TTTCAAGATTTGCTGACGAGTCTCAAGACAGACGACGACCTTCAACAAATGCCA 2163
Db 2130 TTCCATGAAATTCGCTTACGAGAGTCTGAGAACAGACGACGCGCTTTACCAAGTACA 2189
Qy 2164 ACCATGGAAGTCTATGAGTGTCCAGTTCAGGTAAATTTGGAATATCAAAAGCGCTGTG 2223
Db 2190 ACCATGGAAGTGTATGAGTGTCCAGAGATCAGGCAAGTGTGATCATTTAAAGCGCACTC 2249
Qy 2224 ACTAAGAAAGATCTGCTGTGATGTCGAAGAGAAACTGCGCAAGATCATCAGGAT 2283
Db 2250 ACCAAAAAGATCTATGAGTGTAGCGCCAGAGAAAGAAATCTGTCAAGAAATTAAGGAC 2309
Qy 2284 GTAAGAGATGAGACGATGATGTTGCTGTAGAGCTGTGCAATTCAGTCTTCTTAAT 2343
Db 2310 GTCAAGAAATGTAAGGCTGAGACGTCAATGCTCAAGAACTGTGACTGATCTTGAAT 2369

OY	2334	GGGGTTAAGCACCCCGTTAAACACTGTGACATTTGAATGAGGCATTTTGCTGCGCATGACAGG	2403
Db	2370	GGATGCAACACCCCGTAGAGACCTGTATATGACGAAGCTTTTGCTGTGATGACAGT	2429
OY	2404	ACGCTGCTGGACACTGATGGCATCGTCAAACTTAAGAAATGATATTTGTGGGGGACCA	2463
Db	2430	ACTCTCAAGGGCTCATAGCCATTATTAAGCCTTAATAAAGCAGTCTCTGGGGGATCC	2489
OY	2464	AAACATGCGGCTTCTTTAACATGATGTGCTGAAAGTACATTTTAACATGACATATGC	2523
Db	2490	AAACGTCGCGGTTTTTTTAAACATGATGTGCTGAAAGTACATTTTAAACAGAGATTTGC	2549
OY	2524	ACTGAAGTGTACCATTAAGAAAGCATCTCTAGAGAGTGCACACAGACTGTAAACCGCATCGTC	2583
Db	2550	ACACAGTCTTCCACAAAAGCATCTCTGCGGTGACCTAAATCTGTACCTTCCGTGCTC	2609
OY	2584	TCACGCTCTTCTACGACAAAGGAAATGAAGCGTTAAACCATGTGCTGATTAATCATC	2643
Db	2610	TCAACCTTGTTTTACGACAAAATAATGAAAGCAGATCCGAAAGACATTAAGATTGTG	2669
OY	2644	ATAGATACCAAGGACCAACAAAGCCGACAAAGATGATCTGATTTCACTGTTTCAGA	2703
Db	2670	ATTGACATACCGGAGTACCAACCTTAAGCAGCATCTCATTTCTACTGTTTTCAGA	2729
OY	2704	GGATGGGTGAACACAGCTACAGTTGACTACAAAATCAAGAAATCATGATGCGGCTGCA	2763
Db	2730	GGGTGGGTGAAGCAGTTGCAATTAATTAACAAGGCAACGAATATATGACGGCAGTGC	2789
OY	2764	TCGCAGAGCTTACCGCGAAAGCGTTATGTCTGTCAAGTACAAAGTCAACGAGATCA	2823
Db	2790	TCTCAAGGGCTGACCCCTTAAGATGTGTATGTCGTTCCGTACAGGTAAATGAATCT	2849
OY	2824	CTCACTGTGACACTTCTGAGACGCTGAACGCTACTTAAACGACAGAAACGCAAT	2883
Db	2850	CTGTACGACCCCACTTCAGAAATGTGAACGTCTACTGACCCGACGGAAGACCGCATC	2909
OY	2884	GTTCTGAAGACGTACGTGTGTGATCCCTGATTAAGACACTTAACGCTAAATATCCGGG	2943
Db	2910	GTGTGGAACAACTACGTCCGGCGACCCCATGATTAATAACACTGACCTGCMAATACCTGGG	2969
OY	2944	GATTTCAAGGCTTCAATTGAGACACTGGAGGCGCAACAGACCATTAATGACACGCTT	3003
Db	2970	AATTTCACTGCGCAGATAGAGAGTGGGAAAGCAGCATGATGATCATATATAGCAGCATC	3029
OY	3004	CTGTATTAAGCCGACAGACAGCTGATGTGTCCAGATTAAGTGAACGTCTGTGGGCGAAG	3063
Db	3030	TTGGAGAACCGGACCTTACCGAGCTTTCAGAAATTAAGCAACGTGTGTGGGCCAAG	3089
OY	3064	GCTTTAGACCACTTTGGCCACGGCCAACTTTGTCTGACGACAGACAGTGGGAGCG	3123
Db	3090	GCTTTAGGCGCGGTCTGAAGACCGCTGGCATAGCATGACCACTGAACATATGMAACT	3149
OY	3124	TTGCACCCATTCAAGCATGACGAGACGTAATCTCACTGAATATGGCACTGAACCTTCTTTGC	3183
Db	3150	GTGATTTATTTTGAACCGACAAAGCTCACTCAGACGAAATATGTTAAACCACTATATGC	3209
OY	3184	ACCAGGTTCTTTGAGTAGACCTTGACAGTGGTATTTTCCGCTCTTACCGCTGCACCTT	3243
Db	3210	GTGAGGTTCTTTGAGCTGATCTGAGACTCCGGTCAATTTTTCGACCCACACTGTTCCGTTA	3269
OY	3244	ACTTAACAGGATACGACCTGGGATTAATCCGACAGGAAAGATCTATATGGCTTAATTGA	3303
Db	3270	TTCATTAAAGAAATATCACTGGGATATATCTCCCGTGCCTTAACATGTACGGGCTGAATAA	3329
OY	3304	GAGGTAGCAAAAGAGTTGTCAACGCGGATATCCGTGATCAACAAAACCGTTGACACAGGC	3363
Db	3330	GAAGTGTCCTGTCAGCTCTCTTCGACGATACCAACATCGCTCGGGGACGTTGCCACTGGA	3389
OY	3364	AGGGTACGTGATTAAGGAATATATCAATCAAGACATCTCTCCACAATTAATGTGTT	3423
Db	3390	AGAGCTATGTACATGACATCTGTACACTGGCATAATTAATCCGCGATTAACCTATGTA	3449

QY	3424	CCATTAAATGCGCGTTGGCCCCACTCGCTGTATCGTTGACCAAGGACAGGGTAACT	3483
Db	3450	CTGTAAACAAAGAACTGCTTCAAGCTTTAGTCTCCACCAATATGAACACCAAGCT	3509
QY	3484	GATCAACAGCGGATTCCTATCTTAATATGAAGGGCAAACTGTGTGGATCGCGATCT	3543
Db	3510	GACTTTTCTTCAATTGCTACGAATAATGAAGGACGAACCTGTCCTGGTGGTGGGAAAG	3569
QY	3544	ATCAGCAATTCAGGAAAGAAAGTAGAGTCATGGGTCCATTGGCCATAATCATCAGG	3603
QY	3570	TTGTCCGCTCCAGGCAAAATGGTTGACTGTTGTCAAGCCGGCTGAGGCTTACCTCAGA	3629
Db	3630	TGTGATCTCGATTTGGAAATACCTAGCCAGTCCGTTAAATATGACATTAATCTTGTCAAT	3663
QY	3664	GTTAGAGACCCCGTACAGGAACATCATCTAACCAAGTCGAGATCAAGCTATCCACCAC	3723
Db	3690	GTGAGAACCCCATTTAAATACATCATCTATACGAGTGTGAAGACATGCACTTAAGCTT	3749
QY	3724	AGCATGCTAACGTSTAAGCTGTCCACACCTGAACACTGGCGGAACATGTGGCTATA	3783
Db	3750	AGCATGTTGACCAAGAAAGCTTGTGCACTGAATCCCGCGGAACCTGTGTCAAGATA	3809
QY	3784	GGGTAATGGGCTTCTGTGATCGCGCAACCGAATAATTCATCACTGGGTGGACGCTCATTT	3843
Db	3810	GGTAATGGTTAAGCTGTACAGGGCCAGGAAAGCATCACTTGGTGTATAAGCCGGCAGTTC	3869
QY	3844	AGGTTAACCCGTCTGTCTGTACGCTTAAGAACCTCCGAAATACTGAAGTTCTCTTGCTG	3903
Db	3870	AAAGTTTTCGGGATGATGCAACCGAAATCTCACTTAAGAGACGAAAGTTCTGTGTTGTA	3929
QY	3904	TTCTTCCGCAAGGACAAACGCGAACCAACATGATACAGGACAGACTCGGTGTAGTCTT	3963
Db	3930	TTCAATGGGTATCGATTCGCAAGGCGCCGTATCGCAATCTTATCAAGCTTTATCAACCTTG	3989
QY	3964	GACAAATCTATCAAGGTCACCAAGTATCGAGCGAGGAGAGCTCCAGCTACAGAGTG	4023
Db	3990	ACCAACATTTTATCAGTTTCCAGTCTCCAGTCTCCAGAACCCGGAATGTGCACTTCATCATGTG	4049
QY	4024	ATCAGAGGTGACATTAAGCAAGAGGCTGTACCAACCTATTCGTTATGCTGCTAAATAGCAAA	4083
Db	4050	GTGCGAGGGGATTTGTCACCGGCCACCGAAGAGTGAATTATTAATGCTGCTTAACAGCAA	4109
QY	4084	GGTCAACAGGTTCCCGAGTGTGCGGTGCACTGTACCGAAATATGGCCGCTGCTTTGAT	4143
Db	4110	GGACAACCTCGCGGAGGGGTGTGGGAGCGCTGTATTAAGAAATTTCCCGGAAGCTTGAT	4169
QY	4144	AGACAGCCAAATAGCTGTGCGGACGGCTTAAGCTTGTGAACGCAAC---GCTATCATTA	4200
Db	4170	TTAAGCGCGAATCGAAGTAGGAAAGCCGCACTGCTCAAAAGGTGCAAGCTTAACATATCATTT	4229
QY	4201	CATGCTGTAGGACCCCAATTTTCTTAAGATGCGGAAACCGAGGGCGACTTAAGCTGCA	4260
Db	4230	CATGCGGTAGGACCAAACTTCAACAAAGTTTCCGAGGTTGAAGGTGCAAAACAATTGGCA	4289
QY	4261	GCTGCTTAATAGACATAGCGTCCATGCTGCACACCTGACGGGATTAACAAAATATTCAGTA	4320
Db	4290	GAGGCTTAATAGTGCATATGCTTAATATTTGTACAGATTAACATTAACAAGTCAGTAGCATTT	4349
QY	4321	CCGCTACTGTCAACCGGCATCTATTTCTGTGGCAAGAAATGCAATGATGCAATCTTCAT	4380
Db	4350	CCACTGTTGTCCACCGGCAATCTTTTCCGGGAACAATAATGCATTAACCAATCATTTGAAC	4409
QY	4381	CACCTGTTCACTGCTTTGACACCTAAGGATGCCATGTCAACCAATATTTGCTTGGATATA	4440
Db	4410	CATTTGCTGACAGCTTTAAGACACATGATGCAATGTATGCAATATCTGACAGGGAACAG	4469
QY	4441	CAATGGGAGACACAGATTAATGAGGCAATTCACCGCAAGAAAGAAACGTCAAAATCTGGA-	4499
Db	4470	AAATGGGAAATGACTCTTCAGGAAGCAAGTGTGAGAGAGAACGATGTGAGAGATATGC	4529
QY	4500	-----TGATGACAAACCAAGTAGACATTTGATCAGGGTTCACCAAC	4545

4530 ATATCCGACGACTCTTCAGTACAGAACTGATGAGAGTGGTGCATCCGAG 4589
4546 AGCTTTGGGAGGAGACAGGTTACTCCGTCATAGAGGCAAGTTATTCATCTG 4605
4590 AGTTCTTTGGTGGAGGAGGCTTACAGCACAGGATGCGCAAACTTTCATATTTG 4649
4606 GAAAGTACACGATTCATCAGACCGCCAGAGCATTCGCAAAATCCATGCAATGGCCC 4665
4650 GAAAGGACCAAGTTTCCACAGCGCCGCAAGATATAGCAAAATTAATGCCATGTGGCCC 4709
4666 AACAAATCTGAGGCTATGACAGATTTGCTTGTATCATCTCGGGGAGAGATGTCCAGC 4725
4710 GTTGCAACGAGGCGCAATGAGACAGATGATGATCTCTCGAGAAAGCATGAGCAGT 4769
4726 ATCCGCTCCAAATGCCGATGAGAGATCAGAGGCTGTCTCCACTCACAACATTCCA 4785
4770 ATTAGGTGAAATGCGCCGCTGGAAGATGGAAGCTCCACACCACTAGACGCTGCT 4829
4786 TGCTGTGTATTAATGAGTATGACGCGCTGAGCGCTATACAGTTGCGCTGCGAAGAA 4845
4830 TGCTTTGTCATTCATGCTCATATCTCCAGAAAGATACGCGCTAAAAGCTCAGCTCA 4889
4846 GAAAGTTCGCGGATGCTCATCTCTGTTGCCAAGTACAGGATCAGAGCGCTGAG 4905
4890 GAAAGTATTAATGCTGCTCATCTCTTCCATTCGCGAAGTATAGATCATCTGTGTGAG 4949
4906 AAGCTACAGTGCAGCAACGAGTCTCTTTTTCAGGCGCTGATACCAAGGCTGTACACCC 4965
4950 AAGATTCAAATGCTCCAGCTTATTTGTTTCAACGAAAGTCCGCTGATATTCATCA 5009
4966 AGGAAGTA 4973
5010 AGGAAGTA 5017

RESULT 3

AAD29131 standard; DNA; 12523 BP.

07-MAY-2002 (first entry)
Venezuelan equine encephalitis virus (VEE) replicon plasmid.
Venezuelan equine encephalitis virus; VEE; vaccine; HIV infection;
human immunodeficiency virus; alpha-virus replicon; immune response;
therapy; ds.
Chimeric - Venezuelan equine encephalitis virus.
Chimeric - Unidentified.
MO20203917-A2.
17-JAN-2002.
09-JUL-2001; 2001WO-US21701.
07-JUL-2000; 2000US-21695P.
(ALPH-) ALPHAVAX INC.
(UNIC-) UNIV NORTH CAROLINA.
Olmedo R, Keith P, Dryga S, Caley I, Maughan M, Johnston R;
Davis N, Swanson R;
WPI; 2002-171664/22.

Composition useful for treating or preventing HIV infections, comprises
two or more isolated nucleic acids encoding env, gag or pol gene
product of HIV or immunogenic fragment of the gene products -

Example 5; Page 141-145; 201p; English.
The invention relates to a composition comprising isolated nucleic acids
encoding env, gag or pol gene product of human immunodeficiency virus or
immunogenic fragment of the gene products. The gag gene product is
modified to inhibit formation of virus-like particles containing gag
gene product and their release from cells, and the pol gene product is
modified to inhibit reverse transcriptase activity. The invention also
relates to a method for producing an alpha-virus replicon particle used
in vaccines. The composition is useful for inducing an immune response
to human immunodeficiency virus (HIV) or for treating or preventing HIV
infection in a subject. The alpha-virus replicon particle is useful in a
vaccine. The composition is useful for administering a protein or peptide
to a subject. A composition comprising heparin affinity-purified alpha-
virus replicon particle is useful as a clinical trial material and as a
commercial product. The present sequence is Venezuelan equine
encephalitis virus (VEE) replicon plasmid containing the Dva22 gag gene
(p3-40.1.6) used in the exemplification of the invention.

Sequence 12523 BP; 3579 A; 2959 C; 3160 G; 2825 T; 0 other;

Query Match 42.3%; Score 2116.4; DB 24; Length 12523;

Best Local Similarity 64.5%; Pred. No. 0;

Matches 3218; Conservative 0; Mismatches 1746; Indels 24; Gaps 3;

10 ACTAATGATCCATATGGAAGAAATTCAGTTGACTAGATGTCAGAGCCGATATGTC 69
30 AATTACCTACCAAAATGAGAAAGTTTACCTTGACATGAGAAAGAGCCATTCCTC 89
70 AAGTGTACAGCGAGCTTTCACAAATTTGATGATGCAAGAGGAGGTCAGTACAAAT 129
90 AGAGCTTTGAGCGGAGCTTCCGACATTTGAGGTAGAGGAGGAGGAGGAGGAGGAGG 149
130 GACCATGCAATGCGAGAGCTTTTCCATGTCGCAACAAAGCTCATTTAGAGGAGGTC 189
150 GACCATGCTAATGCGAGAGCTTTTCCATGTCGCAACAAAGCTCATTTAGAGGAGGTC 209
190 GACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 249
210 GACCATGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 269
250 CACCGCTATCATTTGATCTGCTATGATGATGATGATGATGATGATGATGATGATGATG 309
270 CACAAGTATCATTTGATCTGCTATGATGATGATGATGATGATGATGATGATGATGATG 329
310 TATGCAAGAGAGCTTTAAGAAAGT-----GACATTACCGAAGAGAGAGAGAGGCTTAA 363
330 TATGCAAGAGAGCTTTAAGAAAGT-----GACATTACCGAAGAGAGAGAGAGGCTTAA 389
364 GCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 423
390 ATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 449
424 CACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 483
450 CACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 509
484 GTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 543
510 GTTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 569
544 ATAGGCTTTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 603
570 ATAGGCTTTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 629
604 AACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 663
630 TCTACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 689
664 GATTTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 723
690 GACGTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 749

QY 724 ACTAATAGATCATATTCTCGGTTGGTTCAACATCTACAGAGAATAGTACTGTGA 783
DB 750 TCCAGCATGTCTATTCTCTGTTGGCTCAGACCATCTACCGAGAGAGGACTTACTG 809
QY 784 CGTAGCTGGCACTTCCAAACGTGTCCACTTGAAGAAAGTCTTAACTTTCACAGGTGA 843
DB 810 AGGAGCTGGGACCTGCGCTGTATTCTACCTTAGGTGCAAGCAAAATTCACATGTGCG 869
QY 844 TGTGGGACCATTTGACCTGTGAAGGATGCTCATCAAAAAGATACAGATCAGCCGAGA 903
DB 870 TGTGAGACTATAGTTAGTTGTGACGGGTACGTCTTAAAGAAATAGCTATCAGTCCAGGC 929
QY 904 CTATACCGTAAAGTTGAGAACTTGGCGTCCCAATGATCGCGAGGTTCTTGAAGTTGC 963
DB 930 CTGATGGGAAGCTTCAAGGCTATGCTGCTACGATGCAACCGCGAGGATTTCTGTCTGC 989
QY 964 AAAGTCAAGATACGCTGCGCGGAGAGGGTTTCTTTGCTGTGTGATGTATGCA 1023
DB 990 AAAGTGACAGACATTTAAACGGGGAGAGGGTCTTTTCCGTTGCAAGTATGTGCA 1049
QY 1024 GCCACACTTTGGCATGACAGAGATTTCTGGCACTGACCTTATGTTGGATGAGCA 1083
DB 1050 GCTACATGTGTGACAAATGACTGGCACTGAGCAAGATGTCAGTGGGAGAGAGCGG 1109
QY 1084 CAAAACTATGTTGGGCTCAACCAAGATGCTCATGTTAGGAGCGCAAAAGAAAT 1143
DB 1110 CAAAACTGCTGTTGGGCTCAACAGGTATGATCTGCAACGGTCCGACCCAGAAAGAAC 1169
QY 1144 ACTAACCAATGCAAACTATCTATTAACAGTGTGCGCCAGGCGTTTTCAGGTGGCG 1203
DB 1170 ACCAATTCATGAAATAATCTTTTGGCCGTAGTGGCCAGGATTTGCTAGGTGGGCA 1229
QY 1204 CGTGAATGCTGCGGCTTGGAGAGAGAAAGACTAGGGGGCGGAGCGGACACTTT 1263
DB 1230 AAGGAATATTAAGAAATCAAGAAAGTAAAGGCCACTAGGACTACAGAGATGACAGTTA 1289
QY 1264 ACTATGGGCTGCTGCTGGGCTTTTCAAGACCCAGAAATTCATCATCTACAAAGGCT 1323
DB 1290 GTATAGGGGTGTGTGTGGCTTTTGAAGGACAAAGATACATCTATTTATTAACGGCCG 1349
QY 1324 GGTAGCAAACTTAAGAAAGTACCTGCGCTTTTGACTGATTTGTGATTCAGGCTT 1383
DB 1350 GATACCCAAACATCATCAAGAGTAAAGACAGCATTTTCCACTATTCGTGCTCCAGAGATA 1409
QY 1384 ACCAGCAACGGGCTCGATATGGGCTTCCGCGGTAGGCTCAAGCTGCTGTGAACCAACT 1443
DB 1410 GGCAGTAACACATTTGAGATGGGCTGAGAAACAAGATCAAGAAATGTTAGAGAGCAC 1469
QY 1444 GTCAAAACCCGACCGGCTATTACATGAGCCGATGTGAGCATCTGCGTGGCTTACAGCA 1503
DB 1470 AAGGAGCGTCACTCTCTCATTTACCGCCGAGAGACGTACAGAAAGCTAAGTGGGACCGAT 1529
QY 1504 GAAGCTGAAGATGGGCTGACGCGGAAGAGATCAGAGAGCCCTGCGCACCTTCTCTCT 1563
DB 1530 GAGGCTAAGGAGGTGCGTGAAGCCGAGAGATTTGCGCGACCTTCAACCTTTTGGCAGCT 1589
QY 1564 GAAATTAAGAAAGAGACCGTATAGAGGAGAAAGTACACCTCATTAAGCAAGAGGAGAGCA 1623
DB 1590 GATGTATGAGGCCACTCTGGAAGCCGATGTCACTTGATGTTACAAAGGCTGAGGCC 1649
QY 1624 GGTAGCTGAGAGACACAGAGACACATCAGGGTGAACAATTACCCAGGCGAAGAGAG 1683
DB 1650 GGCTCAATGGAGACACTCGTGGCTTGATTAAGGTATCAAGCTACGCTGGGAGAGACAG 1709
QY 1684 ATTGGGTCTTAAGCTTATCTTTCAACCCAGCGGATTTGAATVAGTAAAACTGGCGTGT 1743
DB 1710 ATCGGCTCTTAAGCTGTGCTTTCTCCGACGCTGACTCAAGAGTGAATAATTAATCTTGC 1769
QY 1744 ATCCACCATTTGGGGAACAAGTACTGTATGACTCAAAAGTATAGGCGGAGAGATAC 1803
DB 1770 ATCCACCTCTCGCTGAACAAGTATGATTAACACCTCTGGCCGAAAGGCGGTTAT 1829
QY 1804 AAAGTGAGCATACCAAGTAAAGTCAATGTATACAGAAAGGAGCGGCGCTCTGTTCAA 1863

DB 1830 GCGGTGAACCATATCATGTTAAAGTGTGTCAGAGAGGACATGCAATACCGTCCAG 1889
QY 1864 GACTTCCAGGCAATTTAGTGAAGCGCTACGATGCTTTTCAACGAGAGGAGTTCTGTAAC 1923
DB 1890 GACTTTCACACTCTGATGTGAAGTGCACCATGTGTATGAACGAACTGTAGTTCTGTAAC 1949
QY 1924 AGATACCTGACCAACATCGCAATCAACGAGAGCGCTTAAACACTGACGAAGAGTACTAT 1983
DB 1950 AGGTACTTGACCATATTTGCTCACATGAGAGAGCGCTGAACACTGATGAAGATATTAC 2009
QY 1984 AAGACTGTAAAGACTCAGACACAGACTCAAAATACGTTCTGATATGACGACGAAAG 2043
DB 2010 AAAACTGTCAAGCCAGCGAGACAGAGCGGAATTAAGTATACGATACGACAGAAACAG 2069
QY 2044 TGTGTTAAGGAGAAAGACGAGGTCCCTTGTGCTTAACCGGTGATCTGTGATCCACA 2103
DB 2070 TGGGTCAAGAAAGAACTAGTCACTGGGCTAGGGCTCAAGGCGAGCTGTGGATCTTCC 2129
QY 2104 TTTTCAAGGTTTGGTACGAGAGTCTCAAGACAGACAGACAGACCTTCAAAAGCTCCA 2163
DB 2130 TTTCTATTAATTCGCTACGAGAGTCTGAGAACAGACACAGCCGCTCTTACAAAGTACA 2189
QY 2164 ACCATCGAGTCTATGAGAGTCCAGGTTCAAGTAAATCTGGAATCATCAAAAGCGCTG 2223
DB 2190 ACCATAGGGGTATGAGCGTGCAGAGATCAGGCAAGTCTGGCATTTAAAGGCGAGTC 2249
QY 2224 ACTTAAGAAATCTGTTGTGATGTGCAAGAAAGAAACCTGCCAAGATCATCAGGAT 2283
DB 2250 ACCAAAAAATCTAGTGTGAGCGCCMAAGAAAACTGTGCAGAAATTAATAAGGAC 2309
QY 2284 GTAAGAGATGAGACGTATGATGTTGCTGCTAGACTGCTGATTCAGTCTTCTTAAT 2343
DB 2310 GTCAAGAAATGAAGAGGCTGAGCTCAATGCCAAGCTGTGACTAGTCTCTTGAAT 2369
QY 2344 GGGGTTAAGCAACCCGTTAAACCTCTGTACATTTGATGAGGCAATTTGCTGCATGAGG 2403
DB 2370 GAGTGAACCAACCCGTAAGAGACCTGTATATTGAACAACTTTTGTGTGATGACAGGT 2429
QY 2404 ACGTCTGCGCATGATTTGCCATCTGTAAACCTTAAGAAAGTGTATGTGCGGGAGCCA 2463
DB 2430 ACTCTCAGACCGCTCATAGGCTATTAAGACCTTAAAGGACAGTCTCTCGGGAGTCC 2489
QY 2464 AAACAAAGCGGCTCTTTTAAATGATGTGCTGAAAGTATTAACATTAACATATAGC 2523
DB 2490 AAACAGTCCGTTTTTTTAACTATGATGTGCTGAAGTGCATTTTAAACAGAGATTTGC 2549
QY 2524 ACTGAAGTATCAATAAAGCATCTCTAGAGGTGACACAGACTGTAAACCGCATCTGC 2583
DB 2550 ACACAAGTCTTCCACAAAGCATCTCTCGCGTTGCACTTAATCTGACTTGGCTGCTC 2609
QY 2584 TCCACGCTCTTCTTCAAGCAAGCCAAATGAAGCGTTAACCCATGTGATTAATATATC 2643
DB 2610 TCAACCTTGTTTTAAAGCAAAATAATGAAGCAAGATCCGAAGAGACTTAAGATTGTG 2669
QY 2644 ATAGATACCAAGGAGCAACAAAGCCGCAAAAGATATGATCTGATTTTAACTCTGTTTCA 2703
DB 2670 ATTGACACTACCGGACATACCAACCTAAGCAGAGCATCTCATCTTCTGTTTCA 2729
QY 2704 GATAGGTGAAGACGCTACAGATTGACTTACAAAATACGAAATCATGACTGCGCTGCA 2763
DB 2730 GGGTGGGTGAAGCAGTTGCAAAATGATTAACAAGGCAACGAAATTAAGAGGCGAGTGC 2789
QY 2764 TCGCAAGACTTACGCGGAAAGCGTTTATGCTGTCAAGGTACAAAGTCAACGAGAAATCA 2823
DB 2790 TCTCAAGGCTGACCCGTTAAAGGTGTATGCGTTGCTGTTCAAGGTGAATGAATAATCT 2849
QY 2824 CTCTACTCGACACTTCTGAGACGTTGAAGCTGTATCTTCAACGACAGAAAAAGCATT 2883
DB 2850 CTGTACGACCCACCTAGAACATGTAAAGTCTTACTGACCGGACGAGAGCCGATC 2909
QY 2884 GTCTGAAGACGCTAGTGTGATCTCTGATTAAGCACTTACAGGTAATATCTCCGG 2943

Db 2910 GTGTGAAAAACATGACCGGCGACCCATGATTAATAACATGACTGCGCAAGTACCTTGGG 2969
 Qy 2944 GATTTCACGGCTTTCATTTGGACACTGCGACGCGGAACAGAGCCATTATGSCAGCGTT 3003
 Db 2970 AATTTCACGGCAGATAGAGAGAGTGGACAGAGACATGATGCCATCATGAGGACATC 3029
 Qy 3004 CTGTGTAAGCCGACAGACAGTGTGTTCAGAAATAGGTGAACGTCCTGCTGGCGAAG 3063
 Db 3030 TTGGAGAACCGGACCTTACCCAGCGTCTCCAGAAATAGGCAAAAGTGTGTTGGCCAAAG 3089
 Qy 3064 GCTTTAGAGCCAGTCTTGGCCACCGCCCAACATTTGTGCTGACGAGACAGAGTGGAGAGC 3123
 Db 3090 GCTTTAGTCCGGTCTGAAGCCCGCTGCGATAGACATGACCACTGAACATAGGAACACT 3149
 Qy 3124 TTGACCCATTCAACATGACAGAGCGTACTCACTGAAATGGCACTGAACTTCTTTGGC 3183
 Db 3150 GTGATTTATTTTGAACGGAACAAAGCTCATCTGACAGAGATGATTTGAAACCACTATGC 3209
 Qy 3184 ACCAGTTCTTTGAGTAGACCTGACAGAGTGGTTATTTCCGCTCTTACCTGCGCACTT 3243
 Db 3210 GTGAGTTCTTTGACTCGATCTGACTCGGCTTATTTTCGACCCCACTGTTCCGTTA 3269
 Qy 3244 ACTTACAGGATCAGACCTGGATTAATCTGCCAGGAAGAACATGTATGGCTTATAGA 3303
 Db 3270 TCCATTAGGAATTAACACTGGGATTAATCCCGTCCCTTAACATGTAGCGGCTGAATAA 3329
 Qy 3304 GAGGTAGCAAGAGATTGTCAACGGCGATATCCGTGCATCAAAAAGCGGTGACAGAGC 3363
 Db 3330 GAAGTGTCCGTCACTCTCTGACAGTACCAACATGCTCGGCGAGTTGCCACTGGA 3389
 Qy 3364 AGGGAGCTGATATTAAGAAATTAATCAATCAAGACTACTCTCAACAATTAATGTGTT 3423
 Db 3390 AGAGCTATGACATGAACACTGTGTACACTGGGCATTTATGATCCGGCATTAACCTAGTA 3449
 Qy 3424 CCATTAAATCCCGGCTGCCCACTGTTGATCTTGACACAAAGACAGGGTCAAACT 3483
 Db 3450 CCTGTAAACAGAAAGCTCCTCATGCTTATGCTTCAACATTAATGAACCCACAGAGT 3509
 Qy 3484 GATCACAGCGATTCCTATCTTAAGATGAAGGCAATCTGTTGTGATAGTGGCGATCCT 3543
 Db 3510 GACTTTTCTTCTATCTGACGAATTTGAAGGCAAACTGCTGTCGGGAAAAAG 3569
 Qy 3544 ATCAGCATTTCCAGGAGAAAGTATGATCAATGGTCCATTTGCCATTAATCACTAGG 3603
 Db 3570 TTGTCCGTCACAGGAAATGTTGATGTTGTCAACCGGCTGAGGCTACCTTCA 3629
 Qy 3604 TGTGATCTCGATTTTGGAAATACCTAGCAATGTCGTAATATGACATTAATCTTTGCAAT 3663
 Db 3630 GCTGGCTGGATTTTAGGCTATCCAGGTATGTGCCAAATATGACATTAATTTGTAAAT 3689
 Qy 3664 GTTAGGACCCGCTACAGAAACCATACCTACCAACAGTGCAGAGATCAAGCTATCCACAC 3723
 Db 3690 GTGAGGACCCCATTAATTAATACATCATACAGAGTGTGAAGACCATGCACTTAAGATT 3749
 Qy 3724 AGCATGCTAACGTGTAAGGCTGTCCACCACTGAAACATGCGGGAACATGTTGCTATA 3783
 Db 3750 AGCATTTGACCAAGAAAGCTTTGTGTGATCTGAATCCCGGCGAACTGTGTGACGATA 3809
 Qy 3784 GGGTATGGGCTTGTGATCGCGCAACGAGAAATATCATCTGCGGTGGCAAGCTCATTT 3843
 Db 3810 GGTATATGTTAGCTGACAGGCGCCAGGAAAGCATTTGGTGTATAGCGGGGAGTTTC 3869
 Qy 3844 AGGTTTACCGGTGTCTGACGCTTAAGAAACCTGCCGAAATATCTGAAGTTTCTTCTG 3903
 Db 3870 AAGTTTTCGCGGTATGCAAAACCGAAATCTCTACTTGAAGAGACGAAAGTTCTGTTTGA 3929
 Qy 3904 TTCTTCCGCAAGACAAACGGAACACACATGACAGACAGATCGGTGTAGTGTCTT 3963
 Db 3930 TTCAATGGGTAGCATCGAAGGCCGTGTGACCAATCTTACAACTTTTCAACACTTG 3989
 Qy 3964 GACAAATCTTATCAAGGCTCAACAGGTACAGGAGGAGGAGAGTCCAGCGTACAGAGTG 4023
 Db 3990 ACCAAATTTATACAGGTTTCCAGACTCCAGAGCGGATGTGCACCTCATATATCATGTG 4049

Qy 4024 ATCAGAGTGACATTAGCAAGAGCGCTGACCAAGCTATGTTATGCTCTAATAGCAAA 4083
 Db 4050 GTGCGAGGGGATATTGCCACGCGCCACGAGAGAGATTTATTAATGCTCTTAACAGCAAA 4109
 Qy 4084 GGTCAACAGGTTCCGAGTGTGCGGTGCACTGTACCGAAATAGCGCGCTCTTTGAT 4143
 Db 4110 GGACAACTTGGCGGAGGGGTGTGCGAGCGCTGTATTAAGAAAGTTCCCGAAAGCTTCGAT 4169
 Qy 4144 AGACGCCAATATGCTGTGCGGACGGCTAGACTTGTGAAGACGAAC--GCTCATCAT 4200
 Db 4170 TTACAGCCGATCGAAGTAGAAGAAAGCGGACGTGTCAAAAGGTGACCTAAATATCAT 4229
 Qy 4201 CATGCTTAGAGACCAATTTTTCTTAAGATGCGCGAAACCGAGGGGCGCTTAAGCTCGCA 4260
 Db 4230 CATGCCGTAAGACCAACTTCAACAAAGTTTGGAGTTGAAGGTGACAAACAGTTGGCA 4289
 Qy 4261 GCTGCTTACATGAGCATAGCTTCATCTGTCAACGCTGAGCGGATTAACAAAATATCAGTA 4320
 Db 4290 GAGGCTTATGAGTGCATGCTAAGATTTGTCAACGATTAACAATTACAGTCAAGTACGATT 4349
 Qy 4321 CCGCTACTGTCAACCGGCATCTATTCTGTGCGAAAGATCGAGTATGCAATCATTTGAT 4380
 Db 4350 CCAGTGTGTCCACCGGATCTTTTCCGGAAACAAAGATGCACTAACCAATCATTTGAAC 4409
 Qy 4381 CACCTGTTACATGCTTTGACACTAGCGGATGCGGATGACCATATATTGCTTGAATAA 4440
 Db 4410 CATTTGCTGACAGCTTTTAGACACATGATGAGATGATGACATATCTGACGGAGCAAG 4469
 Qy 4441 CAATGGAGACAGGATTAATGAGGCGCTTACCGCAAGAAAGCGTGAATTTCTGA- 4499
 Db 4470 AATGAGAAATGACTCTCAAGGAAGCAAGTGGTGAAGAGAAAGTGAAGAGATATGC 4529
 Qy 4500 -----TGATGACAGCCAGTATGACATTGACTTGTGAGGTTCCACCAAC 4545
 Db 4530 ATATCCGACGACTCTTCACTGATACAGAACCTGATGACAGTGTGATGAGGTGCATCGAAG 4589
 Qy 4546 AGCTCTTTGGAGGAGACAGGTTACTCGTCAATGAGGCAANTTTATTCATACCTG 4605
 Db 4590 AGTTCTTTGGTGAAGGAGGCTACAGCAACAGGATGGCAAACTTTCATATTTTG 4649
 Qy 4606 GAAAGTACAGCATTCATCAGACCGCCCAAGACATTTGCCAAATCATGCAATGTGGCC 4665
 Db 4650 GAAAGGACCAAGTTTCAACAGGCGCCCAAGATATAGAGAAATTAATGCCATGTGGCC 4709
 Qy 4666 AACAAATGAGGCTATAGACAGATTGCTTTGATCATCTGCGGAGAGATGTCCAGC 4725
 Db 4710 GTTGCAACGAGGCGCAATGACAGGTATGATGTATATCTCGGAAAGACATGAGCAGT 4769
 Qy 4726 ATCCGCTCCAAATGCCAGTATGAGAGATCAAGGCGTGTCTCCACTCACACATTTCCA 4785
 Db 4770 ATTAGTGTGAATGCTCCGCTGTGAAGAGTGGAAAGCTTCAACACCTTAGCAGTGTCT 4829
 Qy 4786 TGCCGTGTAAATTAAGCTATAGCGGTGAGGCGGTATACAGTTTCCGCTTGCAGAAAA 4845
 Db 4830 TGCCTTGTGATCCATGCTCATATCTTCAGAAAGATGACGCGCTTAAGGCTCACGTCCA 4889
 Qy 4846 GAAAGTTCGCGGTATGCTCATCTCTGTTGCCGAAGTACAGATACAGGCGGTGAG 4905
 Db 4890 GAACAAAATTAATGTTGTGCTCATCTTTTCATTTGCCGAAGTATGAATCATCTGTGAG 4949
 Qy 4906 AAGCTACAGTGCACAAACAGTCTCTGTTTACAGCGCTGTATCCACCGGTGTACACCC 4965
 Db 4950 AAGATCCAAATGCTCCAGCTATATTTGTTTCCACGAAAGTGCCTGTATATTCATCA 5009
 Qy 4966 AGGAAGTA 4973
 Db 5010 AGGAAGTA 5017

RESULT 4
 AAD29139
 ID AAD29139 standard; DNA; 12379 BP.

XX	AA029139;
AC	ADT
DT	07-MAY-2002 (first entry)
XX	
DE	Human immunodeficiency virus type 1 (HIV-1) pol plasmid.
KW	Human immunodeficiency virus; vaccine; HIV infection; immune response;
KM	alpha-virus replicon; therapy; pol gene; ds.
OS	Chimeric - Human immunodeficiency virus type 1.
XX	Chimeric - Unidentified.
PN	WO200203917-A2.
PD	17-JAN-2002.
PE	09-JUL-2001; 2001MO-US21701.
XX	07-JUL-2000; 2000US-216995P.
PA	(ALPH-) ALPHAVAX INC.
PA	(UYNC-) UNIV NORTH CAROLINA.
P1	Olmsted R, Keith P, Dryga S, Caley I, Naughtan M, Johnston R;
P1	Davis N, Swanstrom R;
DR	WPI; 2002-171664/22.
XX	
PT	Composition useful for treating or preventing HIV infections, comprises
PT	two or more isolated nucleic acids encoding env, gag or pol gene
FR	product of HIV or immunogenic fragment of the gene products -
PS	Example 5; Page 183-187; 201pp; English.
CC	The invention relates to a composition comprising isolated nucleic acids
CC	encoding env, gag or pol gene product of human immunodeficiency virus or
CC	immunogenic fragment of the gene products. The gag gene product is
CC	modified to inhibit formation of virus-like particles containing gag
CC	gene product and their release from cells, and the pol gene product is
CC	modified to inhibit reverse transcriptase activity. The invention also
CC	relates to a method for producing an alpha-virus replicon particle used
CC	in vaccines. The composition is useful for inducing an immune response
CC	to human immunodeficiency virus (HIV) or for treating or preventing HIV
CC	infection in a subject. The alpha-virus replicon particle is useful in a
CC	vaccine. The composition is useful for administering a protein or peptide
CC	to a subject. A composition comprising heparin affinity-purified alpha-
CC	virus replicon particle is useful as a clinical trial material and as a
CC	commercial product. The present sequence is HIV-1 pol plasmid.
XX	
SQ	Sequence 12379 BP; 3562 A; 2920 C; 3084 G; 2813 T; 0 other;
Query Match	42.3%; Score 2113.2; DB 24; Length 12379;
Best Local Similarity	64.5%; Pred. No. 0;
Matches 3216; Conservative	0; Mismatches 1748; Indels 24; Gaps 3
QY	10 ACTAATCATGCATATGGAAGAAATTACGTTGACTTAAGATGTGACAGCCCGATATC 69
DB	30 AATTACCTAACCCAATAATGAGAATAATTCACTTGACATTCGAGGAAGACGCCCAATTCCTC 89
QY	70 AAGTCGTTACAGCGGAGCGTTTCACAAATTGGATGGAAGCAGGCGAGTCACTGACAAT 129
DB	90 AGAGCTTTGACAGCGGAGCGTTCGCCGAGTTTGAGTGAAGACCAAGCAGGTCACTGATAT 149
QY	130 GACCATGCGCATGCGAGAGCGTTTTCGCATGTGGCAAAGAGCTAATTGAGGCGAAGTC 189
DB	150 GACCATGTGAATGCCAGAGCGGTTTTCGATCTGCGCTTCAAACTAATGAAAACGAGGTG 209
QY	190 GACCGGAGACCAAGTATCTTGGACATTGGAAATGGCCCGTCAAGCATGACATTCCAAT 249
DB	210 GACCATTCGACACGATCTTGAACATTGGAAATGGCCCCGCCCGCAGAAATGATTTCTAAG 269
QY	250 CACCGCTATCATTTGATCTGCCCTTATGATMAAGCGGTGAAGACCCGAGACAGACTACACGG 309

D	b	270	CACAACTTATCATTTGATCTGTCTCCATGAGATGTGCGGAAGATCCGGACAGATTGTATTAAG	329
O	y	310	TATGCAAGAAAGACTTTAAGAAAAAGT-----GACATTACCGACAAGAACTATGCTTTAAG	363
D	b	330	TATGCAACTAAGCTGAGAAAACTGTAAAGAAATTAATGATTAAGGAATTGGACAAGAAA	389
O	y	364	GCGGAGAGCTCTGTGGAAATGATGTCCAAACCAAGACGAGACATCTCAATCTCTGTGTAG	423
D	b	390	ATGAAGGAGCTGCGCCGCTCATGAGCGACCTGTGACTGGAAACTGTAGACTATGTGCTTC	449
O	y	424	CACACAGACGCCAGTGTAGTACTTTGGAAAGTGTAGAGCTATTCACAAATGTGTACGCA	483
D	b	450	CACGACGACGAGTCCGTGTCCGTACCAAGGGCAAGTCGCTGTTTACCAGATGTATACGG	509
O	y	484	GTCCATGCAACCGACATCAATTCATCCACAGAGCGCTTAAAGAGTTAAGCAATTATCTGG	543
D	b	510	GTTGACGAGACCGAACAAGTCTATACCAAGCCAAATTAAGGAGTTAAGATCCGCTACTGG	569
O	y	544	ATAGGCTTTGACAGCAACCCCTTTTATGTACAAAAACATGGCAGTTCTTACCTACTTAC	603
D	b	570	ATAGGCTTTGACACCAACCCCTTTTATGTTAAAGACTTGGCTGAGCATATATCATATAC	629
O	y	604	AACACGAATCGGCGCTGACGAGAGATATTGGAAGACGTAATGTGGCCTCGTAACTCA	663
D	b	630	TCTACCAACTGGGCGGACGAAACCGTGTAAACGGCTCGTAACATAGGCTTATGCACTCT	689
O	y	664	GATCTTCAGAGAGAGCGCTTGGAAAACTCTCAATCTTTAGAGAGAGAGGCTCCACCT	723
D	b	690	GACGTTATGAGAGCGGTCAAGTAGAGAGATGTCAATCTTTAGAAAAGATTTGAAACCA	749
O	y	724	ACTAATTAAGATCATATTTCTCGGTGTTCACAAATCTACACAGAAAGATGATCACTGTTA	783
D	b	750	TCCAACAATGTCTATTTCTCTGTTGGCTGACCATCTACCAAGAGAGGAGACTTACCTG	809
O	y	784	CGTACTGGCANTCTTCCAAAGGTGTCCACTTGAAGAGAAAGTCTTACTTCAAGGTAGA	843
D	b	810	AGAGCTGACACTGCGCTGCTGTATTTCACTTACGTGGCAACAAATTAACACATGTCGG	869
O	y	844	TGTGGAGCAATTTGTAGCTGTGAAAGGATACGTATCAAAAAAGATTAACGATACGCCAGGA	903
D	b	870	TGTGAGACTATATGTTAGTTTGGCAGGGGTACGTCGTTAAAGAAATGCTATGCTCAGCG	929
O	y	904	CTATACGGTAAAGTTGAAGACTTGGCGCTCCACATGCAATGCAATCGGAGGTTTCTTGACTGC	963
D	b	930	CTGTATGGGAAAGCTTTCAGGCTATCTGTCTAGATGACCCGAGAGGATTTCTTGTCCTGC	989
O	y	964	AAAGTCACAGATACCGCTGCGCGGCGAGAGGGTTCTTTTGTCTGTGTACGTATGTACGA	102
D	b	990	AAAGTCACAGACACTTCAACGGGAGAGGGTCTCTTTTCCGTGTGACGTATGTGCGCA	104
O	y	1024	GCCACACTTTTGCATCTCAGATGACAGGAAATTCGAGCAACGTGATAGTGTGATGACGA	108
D	b	1050	GCTACATTTGTGTGACCAATATGCTGGCATCTGGCAACAGATGTCAGTGCGGACACGCG	110
O	y	1084	CAAAAACTATTGTGTGGCTCAACCAAAAGATTGTCTCAATGTGTAGACGCAAGAAAT	114
D	b	1110	CAAAAACTGTGTGTGGCTCAACCAAGGTATATGTCTCAACGGTGTGCAACCCAGAGAAAC	116
O	y	1144	ACTAACACATGTGACGAATCTATTTACCATGTGTGTGCGCCAGGCGTTTTCAGGTGGGG	120
D	b	1170	ACCAATACCATTGAAAAATTAACCTTTTGGCCCGTATGTGGCCCAAGCAATTTGCTAGTGGGCA	122
O	y	1204	CGTGAACATCGTGGCGCACTTGGACGACGAAAGAACTTAGGGGTGGGGAGGCACTTT	126
D	b	1230	AAAGATATTAAGGAAGATTCAGAGATGAAGGCCACTTAGACTACGAGATGTACAGTTA	128
O	y	1264	ACTATGGGCTGTCTGTGGCTTTTCAAGACCCAGAAATTCACATCTTACAGAAAGCCT	132
D	b	1290	GTCAATGGGGTGTGTGTGGCTTTTGAAGGCAACAAGATTAATTAATTAAGCGCCCG	134
O	y	1324	GGTACGCAACATTTAAGAAAGTACCTGCGTCTTTGACTCATTTGTGTATTTCAAGCCTT	138

Dh 1350 GATACCCAAACCATCATCAAAAGTGAACAGCATTTCCACTCATTCGTGCTGCCAGATA 1409
Qy 1384 ACCAGCCAGGGGCTCGATATGCGCTTCCGCGTAGGCTCAAGCTGCTGTTGAACCAACT 1443
Db 1410 GGCACTTAACACATTGGAGATCGGGCTGAAGAACAGAAATCGAATAATGTTAAGAGACAC 1469
Qy 1444 GTCAAAACCCGACCCGGCTATTACAAATGCGCGATGTGAGCATCTGCGTGGCTTACAGCAA 1503
Db 1470 AAGAGCCGTCACCTCTCATTTACCGCCGAGAGATCAAGAAAGCTTAAGTGGCCAGCCAT 1529
Qy 1504 GAAGCTGAAGAGTGGCTGACAGCGGAAGATCAAGAAAGCCCTGCCACTTCTGCTCCT 1563
Db 1530 GAGCGTGAAGAGTGGCTGAAGCCGAGAGTGGCGGACCTCACTCACTTTGGCAGCT 1589
Qy 1584 GAATATAGAAAAAGAACCGTATAGGCGAGAAATAGACTTATGTAAGAGGCGAGAGCA 1623
Db 1590 GATGTTGAGAGGCCCATCTGTGAAGCCGATGTGCACTTATGTTACAGAGGCTGGGACC 1649
Qy 1624 GGTAACGTTGAGACACACGAGAGACATCAGGGTGAACAAGTTACCCAGCGAGAGAGAG 1683
Db 1650 GGCTCAGTGAAGACACTGTGGCTTGAATAAAGTTACCAAGCTACGATGGCGAGACAG 1709
Qy 1684 ATTGGGTCTTAAGCTTAATCTTCAACCCGCGGTATTAATATGTAAGTAAATCGCGCT 1743
Db 1710 ATCGGCTCTTAAGCTGTGCTTCTCCGAGGCTGATCAAGAGTGAATAATATCTTGC 1769
Qy 1744 ATCCACCATTTGGCGGAAACAAGTACTGTAATGATCACTCAAAAGTAAAGGCGAGAGATAC 1803
Db 1770 ATCCACCTCTCTGCTGAACAGTCAATAGATTAACACTCTGGCCGAAAGGGCGTTAT 1829
Qy 1804 AAAGTCAGCCATACACCGGTAAAGTCAATTTGACAGAAAGGAGCGGCTCCCTGTTCA 1863
Db 1830 GCCGTGAACCATACATGTTAAAGTATGATGTCAGAGGAGCATGCAATACCGTCCAG 1889
Qy 1864 GACTTCCAGGCACTTAAGTGAAGCGCTACGATCGTTTCAACGAGAGGAGTTGCTAAAC 1923
Db 1890 GACTTTCAGGCTCTGAGTGAAGGTGCCACATTTGTATCAACGAAAGTGAATCTGTTAAC 1949
Qy 1924 AGATACCTGACCAATGSCAATCAACGAGAGGAGGCTTAACATGACGAAAGATCAAT 1983
Db 1950 AGGTACCTGACCAATGTTGCCACACATGAGAGGCGCTGAACATGATGAATAATATTAC 2009
Qy 1984 AAGATGTAAGACTCAGAGACACAGACTGAAATAGTCTTCGATATTGACCAAGAAAG 2043
Db 2010 AAAAATGTCAGGCCAGAGAGACGAGCGGAAATCTGTACGACATGACAGGAAACAG 2069
Qy 2044 TGTGTTAAGCGAAGAGCGAGGCTCTTGTGCTTACCGGTATGCTGTAGATCCACCA 2103
Db 2070 TGCGTCAAGAAAGAACTAGTCACTGGGCTAGGGCTCACAGCGAGCTGGTGTCTCC 2129
Qy 2104 TTTCAAGATTGGCGTAAAGAGTCTCAAGACACGACCGAGCACTTCAAAATGCCCA 2163
Db 2130 TTCATGAATTCGCTACAGAGTCTGAAGACACAGCCGCTCTTACCAAGTACCA 2189
Qy 2164 ACCATCGAGTCTATGAGTGCAGGTTCAAGTAAATCTGAATCATCAAAAGCCCTGTG 2223
Db 2190 ACCAATAGGGGTATAGCGGTGCAGAGTGAAGCAAGTCTGGCATATTAAGAGCCGAGTC 2249
Qy 2224 ACTAAGAAAGATCTGTTGTGTGAGTGCAGAGAAAGAAACTGCGCAAGAAATCATCAGGAT 2283
Db 2250 ACCAAAAAGATCTAGTGTGAGCGCCCAAGAAAGAAACTGTGCAAGAAATTTAAGGGGAC 2309
Qy 2284 GTAAGAGAGTGAACGTATGATGTTGTGCTAGAGCTGTGATTCAGTCTTCTTAAT 2343
Db 2310 GTCAAGAAATTAAGGGGCTGAGCGTCAATGCAAGACGTGAGTCTCTTTGAT 2369
Qy 2344 GGGGTTAAGCAACCCGTTAACTCTGTACATTTGATGAGGCAATTTGGCTGCATGAGG 2403
Db 2370 GGATCAAAACACCCGTTAGAGACCTGTATATTGAAGAGCTTTTGCTTGTATGACAGT 2429
Qy 2404 ACGCTCTGGAAGTATGCTCATGCTCAAACTTAAGAAAGTGTATTTGTGCGGAGACCA 2463
Db 2430 ACTTCAGAGCGCTCATAGCTATTAAGACCTAAAGAGGAGTGTCTTGGCGGAGTCCC 2489

Qy 2464 AAACATGCGGCTCTTTAAACATGATGTCCTGAAGATACATTTTAACCATGACATATGC 2523
Db 2490 AAACATGCGGTTTTTTTAACATGATGTCCTGAAGATGATTTTAACCATGACATATGC 2549
Qy 2524 ACTGAAGTACCATTAAGATCTCTAGAGAGTGCACACAGACTGTAAACCGCATCTGTC 2583
Db 2550 ACACAAGTCTTCCACAAAGGATCTCTCGCGTTCACATTAATCTGTGACTTCGGTCGT 2609
Qy 2584 TCCAGGCTCTTACAGACAGGAATGAAGACGTTAACCCATGTGCTGATTAATCATC 2643
Db 2610 TCACCTGTTTAAAGAAATAAGAAAGAGAAATCCGAAAGACATGAATTTG 2669
Qy 2644 ATAGATACCAACGAGACCAAAAGCCGACAAAGATGATCTGATTTCACTGTTTCAGA 2703
Db 2670 ATTGACACTACCGAGATACCAAACTTAAGAGAGATCTGATCTGCTGTTTCAGA 2729
Qy 2704 GGATGGGTGAACAGCTACAGTTGACTACAAATAACGAAATCATGACTGCGCTGCA 2763
Db 2730 GGGTGGGTGAAGCAGTTCAATTAATTAACAAAGGCAAGAAATATGACGCGAGCTGCC 2789
Qy 2764 TCGCAAGACTTAACCGGAAAGGCGTTATGCTGACAGTACAAAGTCAAGAGATCCA 2823
Db 2790 TCTCAAGGAGCTGACCGTAAAGTGTGATGCCGTTGCGTACAGGTGAATGAATCCT 2849
Qy 2824 CTCTACGACAGACTTCTGAGCAGTGAACGTTACTTACAGGACAGAAAGACGATT 2883
Db 2850 CTGTACGACCCACTTCAGAAATGTGAAGTCTTACTGACCCGACGAGAGACCGCATC 2909
Qy 2884 GTCTGAAGACGCTAGCTGTGTATCCCTGATTAAGACATTAACGCTTAATATCCCGG 2943
Db 2910 GTGTGAAGAAACATGACCGGAGACCATGAGTAAACATCTGACCGGACGAGTACCTGG 2969
Qy 2944 GATTTACGCGCTTCAATGAGCATGCGACCGGCAACAGAGCCATTAATGCGAGCTT 3003
Db 2970 AATTTCACTGACCATGAGAGAGTGCAGAGCAGATGATGCCATCATGAGGACATC 3029
Qy 3004 CTGTAAAGCCGACAGACGCTGATGTTGTTCCGAATTAAGTGAAGTCTGCGGGCGAAG 3063
Db 3030 TTGAGAGACCGGACCTTACCGACGCTTCCGAATTAAGCAAGTGTGTTGGCGAAG 3089
Qy 3064 GCTTTAGACGAGTCTTGCCCAACGCGCAACATTTGCTGACGAGACAGCAGTGGAGAGC 3123
Db 3090 GCTTTAGTCCGCTGTAAAGCCGCTGACATGACATGACCACTGAACATAGGAACACT 3149
Qy 3124 TTGACCAATTCAGATGACAGAGCTGACCTGAATGCACTGAATCTTTTGC 3183
Db 3150 GTGATTTATTTGAAACGAGCAAGCTCACTCAGCAGAGATGATTTGAACCAACTATGC 3209
Qy 3184 ACCAGTCTTTGAGTGAACCTGGAACGTGGGTTATTTCCGCTCTACCGTGCACATT 3243
Db 3210 GTGAGTCTTTGAGCTGATGATGSACTCCGCTTATTTTTCGACCCACTGTCCTTTA 3269
Qy 3244 ACTTACAGGATTCAGACATGAGATTAATCTGCGAGGAAAGAACATGATATGGCTTAATGA 3303
Db 3270 TCCATTAAGAAATTAATCATGAGGATTAATCCCGTGGCTTAATGATGAGGCTGAATAA 3329
Qy 3304 GAGGTGAAGAGAGTGTCAAGCGGATATCCGTGATCACAAAAGCGTTGACACAGGC 3363
Db 3330 GAAGTGTTCGCTCAGCTCTCTGCAAGTACCAACATGCTCGGCGAGTTGCCACTGGA 3389
Qy 3364 AGGAGTGAATTAAGAAATTAATCAATCAAGACTACTCCCAAAATTAATGTGTT 3423
Db 3390 AGAGTCTATGACATGAACATGTGTAACCTGCGCAATTAATGATCCGCGCTAAACCTAGTA 3449
Qy 3424 CCATTAATTCGCGGTTGCCCATCTGTTGATCGTTGACCAAAAGGACAGGGTCAACT 3483
Db 3450 CCTGTAACAGAAAGCTGCTCATGCTTTAGTCTCCACCAATTAAGAACACCAAGAGT 3509
Qy 3484 GATCACAGCGATTCCTATCTTAAGTGAAGGCAATCTGTGTTGTATGCGGATCTT 3543
Db 3510 GACTTTCTTATTCGTGACGAAATTAAGAGGACAGAACTGCTGCTGCGGAGAAAG 3569

```

Qy 3544 ATCAGCATTCAGGGAGAGAAAGTAGAGTCATGGGTCCATTGCCACTAATACATCAGG 3603
Db 3570 TTGTCGCTCCAGGCAAAATGGTTGACTGTTGTTCAGACGGGCTGAGGCTACCTCAGA 3629
Qy 3604 TGTGATCTCGATTTGGGAATACCTAGCCATGTCCGTAATATGACATTAATCTTTGTCAAT 3663
Db 3630 GCTCGGCTGGATTTAGGATCCAGGTGATGTGCCAAATATGACATTAATTTTGTTCAT 3689
Qy 3664 GTTAGAGCCCCGTCAGGAACCATCACTACCAAGTCGAGATCAAGCATCCACAC 3723
Db 3690 GTGAGAACCCCATTAATTAATACCATCACTATCAGCAGTGTGAAGACCATGCTTAAGCTT 3749
Qy 3724 AGCATGCTAAAGCTGTAAAGCTGTCCACCCTGAACACTGCGGAACAATGTGTGCTATA 3783
Db 3750 AGCATGTTGACCAAGAAAGCTGTCTGTGATCTGAATCCGGCGGAACCTGTGTAGAGTA 3809
Qy 3784 GGGTATGGGCTTGTCTGATCGCGCAACGAAATATCATCTGCGGTGGCAGCTCATTT 3843
Db 3810 GGTATGTTAGCTGTCAGAGGGCCAGCAAGACATCATTTGCTATATGCGCGCAGTTTC 3869
Qy 3844 AGGTTTACCCGTCGTCTGTCAGCTTAAGAACCTGCCGAATAATCTGAGGTTCTCTGCTG 3903
Db 3870 AAGTTTTCGGGGATGATGCAACCCGAATCTCACTTGAAGACGGAAGTTCTGTGTTGTA 3929
Qy 3904 TTCTTCGCAAGACCAACCGCAACCAACATGACAGACAGACTCGTGTAGTGTCTT 3963
Db 3930 TTCAATGGGTACGATTCGCAAGCCCGTACGCAAACTTTACAAAGCTTTATCAACCTTG 3989
Qy 3964 GACACATCTATCAAGGTCACCAAGGTACGAGGAGGAGAGCTCCAGGCTACAGATG 4023
Db 3990 ACCAAACATTTATACAGGTTCCAGACTCCACGAAGCCGATGTGCACTCATATCATGTG 4049
Qy 4024 ATCAGAGGTGACATTAGCAAGAGCGCTGACCAAGCTATCGTTAATGCTGTAATAGCAA 4083
Db 4050 GTGCGAGGGGATTTGTCACGGCCACGAGAGATGATTATTAATGCTGCTACACAGCAA 4109
Qy 4084 GGTCAACACGAGTTCCGAGTGTGCGGTGCTGTAACCGAAATGAGCCGCTGCTTTGAT 4143
Db 4110 GGACACACTGCGGAGGGGTGTGGAGCGCTGTATTAAGAAATTCGCGGAAGAGTTGAT 4169
Qy 4144 AGACAGCCATAGCTGTGCGGACGCTAGACTTGTGAAGACGAAC--GCTCATCTATA 4200
Db 4170 TTACAGCGCATCGAAGTAGAAGAAAGCGGACTGTGCTAAAGGTGACGCTAAACATATCAT 4229
Qy 4201 CATGCTGTAGGACCCCAATTTTCTAAGATGCGGAGACCGGAGGCGACTTAAGCTGCA 4260
Db 4230 CATGCCGATGAGCAAACTTCAACAAAGTTTCGAGGTTGAAGTAGAACAAACAGTTGCA 4289
Qy 4261 GCTGCTTACATGACATAGCTCATGCTCAACGCTGACGCGATTCAGAAATATCACTA 4320
Db 4290 GAGGCTTATGATGTCATGCTGTAAGTTGTCAACGATTAATTAAGTCACTGAGCGATT 4349
Qy 4321 CCGCTACTGTCAACCGGATTTATTTGTGTGGCAAAATGAGATGATCAATCTTTCAT 4380
Db 4350 CCACGTGTGTCACCGGATTTTTCGCGGAACAAAGATCACTAAACCAATCAATTGAAAC 4409
Qy 4381 CACCTGTTCACTGCTTTCGACACTACGAGTCCGATGCTCAATATATTTGTTGATATA 4440
Db 4410 CATTTGCTGACAGCTTTTAGACACCACTGATCAGAGCTGTGAGGGTGCATCCGAAG 4469
Qy 4441 CAATGGAGAACAGAGATATCGAGGCCATTCACCGCAAGAAAGCGTCGAATTTCTGA- 4499
Db 4470 AAATGGGAATGACTCTCAAGGAAGAGCTGTGAGAGAGAGAGAGAGATATGTC 4529
Qy 4500 -----TGATGACAAAGCCAGTAGACATTTGCTGTGAGGGTCCACCCAAAC 4545
Db 4530 ATATCCAGCACTTTTCAGTACAGAACTGATCAGAGCTGTGAGGGTGCATCCGAAG 4589
Qy 4566 AGCTTTTGGAGGACAGACGAGTTATCTCCCTCAATAGGCGCAAGTTGATTTACTG 4605
Db 4590 AGTTCTTTGCTGGAAGAGAGGGCTACGACCAAGCGGTGCGCAAACTTTCTCATATTTG 4649
Qy 4606 GAAAGTACAGATTCATCAGACGCGCAAGACATTTGCCAAATTCATGCAATGTGCGCC 4665

```

```

Db 4650 GAAAGGACCAAGTTTCAACAGGCGGCCAAGATATACAGAAATTAATGCAATGTGCGCC 4709
Qy 4666 AACAATCTGAGGTTATGACAGATTTGCTGTACATCTCTGGGGAGAGATATGCGAC 4725
Db 4710 GTTGCAACGAGGCGCAATGAGCAGATGATGATATCTCTGGAAGAACATGACAGAT 4769
Qy 4726 ATCCGCTCCAAATCCAGTAGAGAGATCAGAGGCGTCTGCTCCACTCACACTTCCA 4785
Db 4770 ATTAGTTCGAATATCCCGCTGGAAGATCGGAAGCTCTCTACACACTTAGACGCTGCT 4829
Qy 4786 TGCTGTGTAATTAACCTATGACGCGCTGAGCGGCTATACAGGTTGCGCTTGCAGAAA 4845
Db 4830 TGCTGTGATCCATTCATGCTATGACTCCAGAAAGATACAGGCGCTTAAAGCTTCAGTCA 4889
Qy 4846 GAAAGTCCCGGATATGCTATCATCTCTGTTGCCGAAGTACAGAGATCAAGGCGTGAG 4905
Db 4890 GAACAAAATTACTGTGTCTATCTCTTTCATGTCGGAAGTATGAAATCACTGGTGTGAG 4949
Qy 4906 AAGCTACAGTGCAGCAACAGCTCTGTTTTCAGGCGCTGACACCGGCTGTACACCCC 4965
Db 4950 AAGATCCAAATGCTCCAGGCTATATTTGTTCTCACCGAAAGTCCGTGATATTCATCA 5009
Qy 4966 AGGAAGTA 4973
Db 5010 AGGAAGTA 5017

RESULT 5
AAD29141
ID AAD29141 standard; DNA; 13584 BP.
XX
AC AAD29141;
XX
DT 07-MAY-2002 (first entry)
XX
DE Venezuelan equine encephalitis virus (VEE) RNA replicon plasmid.
XX
KW Venezuelan equine encephalitis virus; VEE; vaccine; HIV infection;
KW human immunodeficiency virus; alpha-virus replicon; immune response;
KW therapy; ds.
XX
OS Chimeric - Venezuelan equine encephalitis virus.
XX
EN WO200203917-A2.
XX
PD 17-JAN-2002.
XX
PF 09-JUL-2001; 2001WO-US21701.
XX
PR 07-JUL-2000; 2000US-216995P.
XX
PA (ALPH-) ALPHAVAX INC.
PA (UNNC-) UNIV NORTH CAROLINA.
XX
PI Olmsted R, Keith P, Dryga S, Caley I, Maughan M, Johnston R;
PI Davis N, Swanstrom R;
XX
DR WPI; 2002-171664/22.
XX
PT Composition useful for treating or preventing HIV infections, comprises
PT two or more isolated nucleic acids encoding env, gag or pol gene
PT product of HIV or immunogenic fragment of the gene products
XX
PS Example 5; Page 190-195; 201pp; English.
XX
CC The invention relates to a composition comprising isolated nucleic acids
CC encoding env, gag or pol gene product of human immunodeficiency virus or
CC immunogenic fragment of the gene products. The gag gene product is
CC modified to inhibit formation of virus-like particles containing gag
CC gene product and their release from cells, and the pol gene product is
CC modified to inhibit reverse transcriptase activity. The invention also

```

relates to a method for producing an alpha-virus replicon particle used in vaccines. The composition is useful for inducing an immune response to human immunodeficiency virus (HIV) or for treating or preventing HIV infection in a subject. The alpha-virus replicon particle is useful in a vaccine. The composition is useful for administering a protein or peptide to a subject. A composition comprising heparin affinity-purified alpha-virus replicon particle is useful as a clinical trial material and as a commercial product. The present sequence is Venezuelan equine encephalitis virus (VEE) RNA replicon plasmid used in the exemplification of the invention.

Sequence 13584 BP; 3942 A; 3114 C; 3392 G; 3136 T; 0 other;

Query Match 42.3%; Score 2113.2; DB 24; Length 13584;
Best Local Similarity 64.5%; Pred. No. 0;
Matches 3216; Conservative 0; Mismatches 1748; Indels 24; Gaps 3;

10 ACTAATGATCCATATATGAAAGAAATTCAGTTGACTAGATGACAGCCGATATGTC 69
Db AATTACCTAACCCAAATGAGAAAGTTCACTGACATGAGAGAGAGAGCCCATTCCTC 89
Qy AAGGTGTAACAGCGGACGTTCCACATTTGATGATGAGAGAGAGAGAGAGAGAGAGAG 129
Db AGAGCTTGACAGCGGAGCTTCCGCAAGTTGAGTAAAGCCAGAGCTCATATAT 149
Qy GACCATGCAATGACAGAGCGTTTTCGATGAGCAACAAAGCTATTGAGAGCGAAGTC 189
Db GACCATGCTAATGACAGAGCGTTTTCGATGAGCAACAAAGCTATTGAGAGAGAGAG 209
Qy GACCGGAGCAAGTTATCTTGACATGAGAGTGGCCGCTGACATGACATTCAT 249
Db GACCATGCGACACCATCTTGACATGAGAGTGGCCGCTGACATGATTTCTAAG 269
Qy CACCGCTATCATTTGATGCTGCTGATGATGAGCGGTGAAGACCCGAGACATTAACAG 309
Db CACAGATATCATTTGATGCTGCTGATGATGAGCGGTGAAGACCCGAGACATTTAAG 329
Qy TATGAGAGAAAGCTTAAGAAAAAGT-----GACATTACGACAAAGATGAGCTCTAAG 363
Db TATGAGAACTAAGCTTAAGAAAAAGT-----GACATTACGACAAAGATGAGCAAGAA 389
Qy GCGGACAGCTGCTGGAAGTATGCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 423
Db ATGAAGAGAGCTGCGCGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 449
Qy CACACAGAGCGCAGCTGATGAGTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 483
Db CACAGAGAGAGCTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 509
Qy GTCCATGACCGAGATCATATCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 543
Db GTTACGAG 569
Qy ATAGGCTTTGACAG 603
Db ATAGGCTTTGACAG 629
Qy AACACGAG 663
Db TCTACCAAG 689
Qy GATCTTACAG 723
Db GACGTTATGAG 749
Qy ACTAATAGATCATATTTCTGCTGCTTCAACATCTACAGAGAGAGAGAGAGAGAGAG 783
Db TCAACATATGTTCTATTTCTGCTGCTTCAACATCTACAGAGAGAGAGAGAGAGAG 809
Qy CGTAGCTGAGATCTTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 843
Db AG 869

844 TGTGGACCAATGTCAGCTGTGAAGGATACATCAAAAAGATACATGAGAGAGAGAG 903
Db TGTGAGACTATAGTATGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 929
Qy CTATACGCTAAAGTTGAG 963
Db CTGATGAG 989
Qy AAGACAG 1023
Db AAGTACAG 1049
Qy GCCACACTTTCGATACAGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1083
Db GCTACATGTTGTGACCAATATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1109
Qy CAAAACCTATGTTGGGCTCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1143
Db CAAAACCTGCTGGGCTCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1169
Qy ACTAACAAG 1203
Db ACCATACCAAG 1229
Qy CGTGAACATCGTGCAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1263
Db AAGGATATTAAG 1289
Qy ACTATGGCTGCTGCTGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1323
Db GTCAATGGGCTGTTGTTGGCTTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1349
Qy GGTACGCAAAATTAAG 1383
Db GATACCCAAACATATCAAG 1409
Qy ACCAGCCAGGCTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1443
Db GGCAGTAAACATGAG 1469
Qy GTCAAAACCGGAG 1503
Db AAGAGCCGTCACCTCTATACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1529
Qy GAACTGAAAG 1563
Db GAGGTAAG 1589
Qy GAAATGAG 1623
Db GATGTTAG 1649
Qy GGTAGCTGAG 1683
Db GGTGAG 1709
Qy ATTGGCTTTACGCTATATCTTCAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1743
Db ATCGGCTTTACGCTATATCTTCAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1769
Qy ATCCACCACTGAG 1803
Db ATCCACCTCTGCTGAG 1829
Qy AAGTGCAG 1863
Db GCCGAG 1889
Qy GACTTTCAG 1923
Db GACTTTCAAG 1949
Qy AGATACGTCAG 1983

1950 AGGACTGACCACTATTTGCCACACATGAGAGACGGCTGAACATGATGAAATATTTAC 2009
1984 AAGACTGAAGACTCAGAGACACAGACTCAGAAATACGTTCTGATATTGACGACGAAAG 2043
2010 AAAAATCTGCAAGCCAGAGACACAGAGCCAAATACCTGTACGACATCGACAGGAAACAG 2069
2044 TGTGTTAAGGAGAAAGACGAGGTCCCTGTGCTTAACCGGTGATCTGTGATGCCACA 2103
2070 TGCCTCAAGAAAGAACTAATGTCATGAGGTAGGGCTCAGAGGCGAGCTGTGGATCTTCC 2129
2104 TTTACAGATTGGCTACGAGGTCTCAAGACAGACGACGAGCACTCAGAAAGTCCCA 2163
2130 TTCCATGATTTCCGCTACGAGAGCTGAGAACAGACGACGCGCTCTTATCAAGTACCA 2189
2164 ACCATCGAGTCTATGAGAGTCCAGGTTCAGTAAATTTGAAATCATCAAAAGCGCTGTG 2223
2190 ACCATAGGGGTGTATGGGTGCTCCAGAGTCAAGCAAGTCTGGCATCATTTAAAGCGCAGTC 2249
2224 ACTAAGAAAGATCTGTTGTGAGTGCAGAAAGGAAAACTGGGAGAAATCATCAGGGAT 2283
2250 ACCAAAAAGATCTAATGTGTAGCGCCAGAAAGAAATCTGCAAGAAATTTAAGGAGAC 2309
2284 GTAAGAGATGAGACGATGATGATGTGCTGTAAGACTGTGATTCAGTCTTTCTAAAT 2343
2310 GTCAAGAAATGAAAGGCTGAGCGTCAATGCCAGAACTGTGACTGACTCTTGAT 2369
2344 GGGGTTAAGACCCCGTTTAACTCTGTACATTTGATGAGCAATTTGCTGCTGACAGGG 2403
2370 GGATGCAAAACACCCCGTATGAGACCCCTGTATTTGACGAAGCTTTTGCTGTACAGAGT 2429
2404 ACGGTGCTGGACATGATGCGATGCTCAACCTTAAGAAAGTGTATGTGGGGGAGCCCA 2463
2430 ACTTTCAGAGGCTCATAGCACTTATTAAGCTTAAAGGACAGTGTCTTGCGGGATCC 2489
2464 AAACATGCGGCTCTTTTAACTATGATGTGCTGAAAGTATTTTAACTATGACATATGC 2523
2490 AAACAGTGGGTTTTTTTAACTATGATGTGCTGAAAGTATTTTAACTATGACATATGC 2549
2524 ACTGAATGTACCAATAAAGATCTTAGAGGTGCACACAGACTGTAAACCGCATGCTC 2583
2550 ACAAGAGTCTTCCAAAGAGCATCTGCGGTGCATTAATCTGTGACTTGGTCTCTC 2609
2584 TCCAGGCTCTCTGACAGAGAAAGTGAAGCGTTAACCATGTGCTGATTAATATCATC 2643
2610 TCAACCTTGTTTACGAAAAAATGAAAGAGATCCGAAAGACATGAAGTGTG 2669
2644 ATAGATACCAAGGACCAAAAGCCGCAAAAGTATGATTTCTAATCTGTTTTCAGA 2703
2670 ATTGACACTACCGGAGTACCAAACTTAAGAGAGAGATCTCATTTCTCACTTGTTCAGA 2729
2704 GGATGGGTGAACAGCTACAGATGTACAAATAACGAATATCATGCTGGCTGCA 2763
2730 GGGGTGGTGAAGCAGTTGCAATAGATTAACAAGCAAGAAATATATACCGCAGCTGCC 2789
2764 TCGCAAGGACTTAGCGGGAAGGGTTTATGCTGCAAGTCAAAAGTGAAGAGAAATCA 2823
2790 TCTCAAGGCTGACCCGTAAGGTGTGTATGCTGCTGTAACAGTGAAGAAATCTT 2849
2824 CTCTACTGACAGCTTCTGACACGTGAACGTGTACTTTAACAAGCAAGAAAAAGCATT 2883
2850 CTGTACGACCACTCAGAACATGTGAACGTCTTACAGACCCGAGAGGAGACGGCATC 2909
2884 GTCTGGAAGACGCTAGCTGTGTATCTCTGATTAAGACATTAACGCTAAATATCTCCGG 2943
2910 GTGTGGAAGAACTAAGCCGCGGACCATGTGATTAAGAAACATGACTGCCAAGATCCCTGG 2969
2944 GATTTCAGGCTTATTTGAGAGACTGAGAGCGGCAACGAGCCATATATGGCAAGCTT 3003
2970 AATTTCATGCAAGATGAGAGTGTGCAAGAGATGATGCCATATATAGGACATC 3029
3004 CTGTATTAAGCCGACAGACTGATGTGTCCAGATTAAGTGAAGCTGTGCTGGCGAAG 3063

3030 TTGAGAGACCGGACCTTACCAGATGTTCCAGATTAAGGCAAACTGTGTGGGCCAAG 3089
3064 GCTTTAAGCAGCTTGTGGCCACGAGCAAGATTGTGCTGACGAGACAGCTGGGAGACG 3123
3090 GCTTTAGTGGCGGTGTAAGACCGCTGAGCATGACATGACACTGAACATATGGAACACT 3149
3124 TTGACCCATTTCAAGCATGACAGAGCTACTCACTGAAATGAGCACTGAACTTTTGTG 3183
3150 GTGATTAATTTTGAAGGACCAAAAGCTCATCTACAGACAGATGATTTGAACCAACTATGC 3209
3184 ACCAGGTTCTTTGAGATGACCTGAGACATGAGTATTTTCGCTCTTACCGTGCATT 3243
3210 GTGAGGTTCTTTGAGACTGATCTGAGCTCGGATCTATTTTCTGACCCACATGTTCCGTTA 3269
3244 ACTTACAGGATCAAGCATGAGGATTAATCCGACGAGGAAAGCAATGTATGGGCTTAATGA 3303
3270 TCCATTGGAATATCACTGGATTAATCCCGTGCCTTAACATGTAAGGGCTGAATAA 3329
3304 GAGGTACAAAGAGTGTACAGCGCATATCCGTGATCACAAAAGCGTTGACACAGGC 3363
3330 GAATGTGCTCGTACGCTCTCTCGAGGTACCCACAACTGCTCGGCGAGTTGCCACTGGA 3389
3364 AGGGTACCTGATATTAAGAAATTAATACATCAAGACTTCTTCCAAATTAATGTGGTT 3423
3390 AAGTCTATGACATGAACACTGTGACACTGCGCAATTAATATCCGCAATPAACCTAGTA 3449
3424 CCATTAATGCGCGTTGCCCACTCGTGTATGCTTGAACCAAAAGACAGGGTACAACT 3483
3450 CTTGTAAACAGAAACTGCTCATGCTTTAGTCTTCCACCAATTAATGAACCCACAGAGT 3509
3484 GATCAGACCGGATCTTATCATAGATGAAGGCAAACTGTGTTGTGATGCGGATCT 3543
3510 GACTTTTCTTATCTGTCAGAAATTAAGGCGAGAACTGTCTGTGTGTGGGAAAG 3569
3544 ATCAGCATTTCCAGGAAGAAATGATAGTCCATGCTCATTTGCCATTAATCATCAGG 3603
3570 TTGTCCGTCCAGGCAAAATGTGTGACTGTGTGCAGACCGGCTGAGGCTTACCTTGA 3629
3604 TGTATCTCAATTTGGGAATACCTAGCAGTGTGCTGAATATGACATTAATCTTGTCAAT 3663
3630 GCTGGCTGATTTAAGCATCTCCAGTGTGTGCCAAATATGACATTAATTTGTAT 3689
3664 GTTAGACCCCGTACAGAAACATCATCAACACAGTGCAGATCAAGCTATCACAC 3723
3690 GTGAGACCCCATTAATATACATCATCATCAGAGTGTGAAGACATGCAATTAAGCTT 3749
3724 AGCATGCTTAACGTGTAAAGGCTGTCCACCACTGAACTGCGGGAACATGTGTGCTATA 3783
3750 AGCATGTTGACAAAGAAAGCTTGTCTGATCTGAATCCGCGGAACCTGTGTACAGATA 3809
3784 GGTATGGGCTTGTGATCGGCAACGGAATATCATCACTGGGTGGACCGCTCATTT 3843
3810 GGTATGTTTACGTGTACAGGCGCAGGAAAGCATATGTGTGTTATGACCGGCAAGTTC 3869
3844 AGTTTACCGGTGTCTGTACAGCTTAAGAACACTGCCGAAATATCTGAGTCTTCTGCTG 3903
3870 AAGTTTCCCGGTATGCAAAACGAAATCTCACTTGAAGAGAGAAAGTTCTGTTTGA 3929
3904 TTTTCCGCAAGAACAGCGCAACCAACATCAATGAGAGAGAGTCCAGCGTATAGTGTCT 3963
3930 TTTCAATGGGTGATGCCAAGGCCGCTGAGCAATCTTTCATTAAGCTTTCATCAACCTTG 3989
3964 GACAACTATTAACAGGTTCAACAGATACAGATACGAGGAGAGAGTCCAGCGTATCAGAGTG 4023
3990 ACCAACTTTATTAACAGGTTCAACCTCAACGAAACCGAGATGCACTCTATATCATGTG 4049
4024 ATCAGAGTGCATTTAGCAAGAGCGCTGACCAAGCTATCGTTATGCTTAATAGCAAA 4083
4050 GTGCGAGGGAATATTTGCCAGCGCAACCGAAGAGATGATTAATATGCTGTCTTAACAGCAA 4109
4084 GGTCAACCAAGGTTCCGAGTGTGCGGTGCATGTATCCGAAATAGCGCGGCTGCTTTGAT 4143
4110 GGACAACTTGCGGAGGGGTGTGCGAGCGCTGTATTAAGAAATTTCCGGAAGACTTCGAT 4169


```

OY 4144 AGACAGCCAAATAGCTGTGGGACGGCTAGACTTGTGAAGCAGAAC---GCTCATCAT 4200
DB 4170 TTACAGCCGATGAGTAGAGAAAAGCGGCACTGTGCAAAAGTGACACTTAATCATTC 4229
OY 4201 CATGCTGATGAGCAACCAATTTTCTTAAGATGCCGGAACCGGAGGCGACTTAAGTCGCA 4260
DB 4230 CATGCCGATGAGCAACCAATTTTCTTAAGATGCCGGAACCGGAGGCGACTTAAGTCGCA 4289
OY 4261 GCTGCTCATGAGATAGCGCTCAATGCTCAAGCGTGAAGCGGATTAACAAAATATCAGTA 4320
DB 4290 GAGGCTTATGATCATGCTAAGATGTGCAACGATTAACATTAACAGTCACTAGCGATT 4349
OY 4321 CCGCTACTGTCAACCGGATCTATTCTGTGGCAAAAGATCGAGTGAATCATTCGAT 4380
DB 4350 CCACTGTGTGCAACCGGATCTTTTCCGGGAACAAAGATCGATTAACCAATCATTTGAAC 4409
OY 4381 CACCTGTTCATGCTTTTGACACTAGCGATGCCGATGTCACCAATATTTGTTGATAA 4440
DB 4410 CATTTGCTGACAGCTTTTGAACACCACTGATGATGATGATTAATGTCAGGAGCAAG 4469
OY 4441 CAATGGAGACCGAGATATGAGGCAATTCACCCCAAGAAAGGTCGAAATTTTGCA- 4499
DB 4470 AAATGGAAATGACTCTCTCAAGAGCAGTGTAGAGAGAAAGCAGTGGAGAGATATGC 4529
OY 4500 -----TGATGACAAAGCCAGTAGACATTGACTTGTCAGGAGTCCACCCAAAC 4545
DB 4530 ATATCCGACGACTCTTCAGTGTACAGAACTGATGACAGAGCTGTGAGGTTGATCCGAG 4589
OY 4546 AGCTTTTGGCAGCAGACCAAGTTACTCTCGTCAATGAGGCAAGTTGATTCATACCTG 4605
DB 4590 AGTTCTTTGGCTGAGAGAGAGGCTACAGACAAACGATGGCAAACTTTTCATATTTG 4649
OY 4606 GAAAGTACACGATTCATCAGACCGCCCAAGGACATTCGCAATTCAGTCAATGGGCC 4665
DB 4650 GAAAGGACCAAGTTTCCACCGGCGCCCAAGGATATAGAGAAATTAATGCCATGTGGCCC 4709
OY 4666 AACCAATTCAGGCTATAGAGAGATTTGCTGTATCATCTCGGAGAGATATGTCCAGC 4725
DB 4710 GTTGCAAGGAGGCCAATGAGAGGATATGATATCTTCGGAGAAAGCATGAGCAGT 4769
OY 4726 ATCCGCTCCAAATGCCAGTAGAGAGATGAGAGGCGCTGTCTCACTCAGACACTTCA 4785
DB 4770 ATTAGTGAATATGCCCGTCCAGAGAGTCCGTAAGGCTCTTCACTAGACGCTGCT 4829
OY 4786 TGCCTGTATTAATGCTATGACGCGCTAGAGCCGCTATACAGTTGCGCTTCCGAGAA 4845
DB 4830 TGCCTGTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4889
OY 4846 GAAAGTTCGCGGATGCTATCATTTCTGTGCGGAAGTACAGATCAGAGGCGTGGAG 4905
DB 4890 GAAAGTATTAATGCTGTGCTATCTTTCATTTGCGGAAGTATGATGATGATGATGATG 4949
OY 4906 AAGCTACAGTGCAGAAACAGTCTGTTTCAGGCGCTGATCAACCGGCTGTACACCC 4965
DB 4950 AAGATTCATGCTCCAGCCTATATTGTTTCAACGGAAGTCCCTGCGATATTCATCA 5009
OY 4966 AGGAAGTA 4973
DB 5010 AGGAAGTA 5017

```

```

RESULT 6
AAD29132
ID AAD29132 standard; DNA; 7479 BP.
XX AAD29132;
XX 07-MAY-2002 (first entry)
DB Venezuelan equine encephalitis virus (VEE) non-structural gene.
XX Venezuelan equine encephalitis virus; VEE; vaccine; HIV infection;
KW

```

```

KW human immunodeficiency virus; alpha-virus replicon; immune response;
KW therapy; non-structural gene; ds.
OS Venezuelan equine encephalitis virus.
XX
XX Key Location/Qualifiers
XX CDS 1..7479
XX /tag= a
XX /product= "VEE non-structural protein"
XX /transl_except= (pos:5635..5643, aa:Gln-Arg)
XX /note= "insertion of 3 bases alters the reading frame,
XX CDS does not include stop codon"
XX /partial
MO200203917-A2.
PD 17-JAN-2002.
PF 09-JUL-2001; 2001WO-US21701.
PR 07-JUL-2000; 2000US-216995P.
PA (ALPH-) ALPHAVAX INC.
PA (UNNC-) UNIV NORTH CAROLINA.
PI Olmsted R, Keith P, Dryga S, Caley I, Maughan M, Johnston R;
PI Davis N, Swanson R;
PI WPI: 2002-171664/22.
PI P-PSDB; AAE18302.
DR Composition useful for treating or preventing HIV infections, comprises
XX two or more isolated nucleic acids encoding env, gag or pol gene
XX product of HIV or immunogenic fragment of the gene products
XX
XX Example 2; Page 145-156; 201pp; English.
XX
XX The invention relates to a composition comprising isolated nucleic acids
XX encoding env, gag or pol gene product of human immunodeficiency virus or
XX immunogenic fragment of the gene products. The gag gene product is
XX modified to inhibit formation of virus-like particles containing gag
XX gene product and their release from cells, and the pol gene product is
XX modified to inhibit reverse transcriptase activity. The invention also
XX relates to a method for producing an alpha-virus replicon particle used
XX in vaccines. The composition is useful for inducing an immune response
XX to human immunodeficiency virus (HIV) or for treating or preventing HIV
XX infection in a subject. The alpha-virus replicon particle is useful in a
XX vaccine. The composition is useful for administering a protein or peptide
XX to a subject. A composition comprising heparin affinity-purified alpha-
XX virus replicon particle is useful as a clinical trial material and as a
XX commercial product. The present sequence is a non-structural gene from
XX Venezuelan equine encephalitis virus (VEE) used in the exemplification
XX of the invention.
SQ Sequence 7479 BP; 2141 A; 1774 C; 1900 G; 1664 T; 0 other;
Query Match 42.2%; Score 2111; DB 24; Length 7479;
Best Local Similarity 64.5%; Freq. No. 0;
Matches 3209; Conservative 0; Mismatches 1740; Indels 24; Gaps 3;
OY 25 ATGGAAGAATTCAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 84
DB 1 ATGGAAGAATTCAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 60
OY 85 AGTTTCCAAATTTGAGATGAGCAAGCAGGCTACTGCAATGACCATGCAATGCC 144
DB 61 AGTTCCCGCAGTTTGAAGTGAAGCCAGGCTACTGATTAATGACCATGCTAATGCC 120
OY 145 AGAGCGTTTGGCATGTGGCAACAAGCTCATTTGAGAGCGGAGTGCAGCCGAGCAATT 204
DB 121 AGAGCGTTTGGCATGTGGCTTCAAAACTGATGCAAAAGGAGGTGAGACCATTCGACACG 180
OY 205 ATCTTGACATTGGAAGTGGCCCGTCAAGATGACATTCATCAATCCGCTATCATGT 264

```

181 ATCTTGACATTTGAAAGTGGCCGCGCGAGAAATGATTTCTTAAGCAACAATATCATTTGT 240
Qy 265 ATTCGCCCTATGATPAAGCGCTGAAGAACCAGACACTCAACGGTATGCAAAAGACTT 324
Db 241 ATCTGCTCGATGAGATGTGCGGAAGATCCGACAGATGTGATTAAGTAACTTAAGCTG 300
Qy 325 AAGAAA-----AGTGAATTACCGACAAAGAACATAGCTCTTAAGGGGGGAGACCTGCTG 378
Db 301 AAGAAAACCTGTAAAGAAATTAATGATTAAGAAATTGGAACAAGAAATGAAGAGCTCGCC 360
Qy 379 GAATGATGTCACACGACGACAGAGATCCATCTCTGTGTATGACACAGACGCGCACG 438
Db 361 GCGGTATGAGCGACCCCTGACCTGGAAACTGAGACTATGTGCTCCACGAGACGAGTGC 420
Qy 439 TGTAGTACTTTGGAAGTGTAGCAGTATCAAGATGTGTAGCGAGTCCATGCAACCGACA 498
Db 421 TGTGCTACGAAGGCAAGTGTGCTGTATCCAGAGATGTATAGCGGTTGACGCGACCA 480
Qy 499 TCAATCTACCAACGAGGCGCTTAAAGAGTTAGGACAATTTACTGGATAGGCTTTGACAG 558
Db 481 AGTCTTATACCAAGCCAAATPAAGGAGTTAGAGTGCCTACTGATAGGCTTTGACACC 540
Qy 559 ACCCTTTTATGTACAAAACATGAGCTTCTACCTTACTTAACAACGAGCTGGGCT. 618
Db 541 ACCCTTTTATGTATTAAGAACTTGGCTGAGACATATCATCTCATCTCAACCACTGGGCC 600
Qy 619 GACGAGAGATTTGGAAGACGTTACCTGGGCTCGTAACTCAGATCTTCAGAGAGC 678
Db 601 GACCAAAACCGTGTAAAGGCTGTAAACATAGGCGCTATGCACTGACGTTATGAGAGCG 660
Qy 679 AGGCTTGGAAAACTCTCAATCCTTAGGAAGAGAGGCTCCAACTACTAATTAAGTACATA 738
Db 661 TCAGCTAGAGGAGATGTCATTTCTTAAGAAAGATTTGAATTAACATCCAACTTTCTTA 720
Qy 739 TTCTCGGTTGTTCAACATCTACACAGAAATGATCACTGTAACTGATGCTGATCTT 798
Db 721 TTCTGTTGGCTGACCATCTACACAGAAAGAGGACTTACTGAGAGAGCTGGGACCTG 780
Qy 799 CCAAAAGCTGTCCACTTGAAGAAAGTCTAACTTACAGATGATGTGGGACCATTTGTC 858
Db 781 CCGTCTGTATTTCACTTACGTGGCAAGCAAAATTAACATGTCCGTGTGACATTAATTT 840
Qy 859 AGCTGTGAAGGTTGCGTATCAAAAAGATTAAGATCAGCCAGGACTTACGGTAAAGTT 918
Db 841 AGTTGCGAGGAGTACGTCTTAAGAAAGATTAAGTCACTCAAGGCTGTATGGAGGCT 900
Qy 919 GAGAACTTGGCGTCCACATGCAATCGCGAGGAGTTCTTGAAGTTGCAAAATCAGATACG 978
Db 901 TCAGGCTATGCTGCTAGCATGCAACGGAGGAGATTTCTGTGCTGAAAGTACAGACACA 960
Qy 979 CTGCGCGCGAGAGGAGTTCTTTTGTGTGTGTAGTATGTAACACAGCCACTTTGGCAT 1038
Db 961 TTAAACGGGAGAGGAGTCTTTTTCCTGTGCAAGTATGTGCAAGCTCACTTGTGTAC 1020
Qy 1039 CAGATGACAGGAGTTTGGCAACTGACGTTAGTGTGATGACGACAAAATCTATTTGGTT 1098
Db 1021 CAAATGATGGCATATCTGGCAACAGATGTCAAGTGTGAGGACGAGCAAAAATCTGTGTT 1080
Qy 1099 GGGGTCAACCAAAAGATTTGCTCAATGTGTAGGACGCAAAAGAAATACTTAACAATGACG 1158
Db 1081 GGGGTCAACCAAGGATATGTGTCAACGCTGCAACCCAGAAACCAACAATACATGAAA 1140
Qy 1159 AACTATCTATTACAGATGTGTGCCAGGCGTTTTCACAGTGGCGCGTGAACATGTGCC 1218
Db 1141 AATTACTTTTGGCCGTGTGTGCGCCAGGCAATTTGTCTAGTGGGCAAAAGAAATATPAAGAA 1200
Qy 1219 GACTTGAACGACGAAGAAAGAACTAGGGGTGTGGGAGGCGCACTCTTACTATGGGCTGCTG 1278
Db 1201 GATCAAGAAAGTGAAGGCGCACTAGACTACAGATTAACAGTTATGTATGGGCTGTGT 1260
Qy 1279 TGGGCTTTCAAGACCCAGAAAATCAATCTTACAGAAAGCTGTGTACCAAACTT 1338

1261 TGGGCTTTTGAAGGCACAAGATPAACATCTATTATTAAGGCGCCGATACCAAAACATC 1320
Qy 1339 AAGAAATACCTGCGCTTTGACTCATTTGTGATTCACAGCCCTTACAGCCAGGGCTC 1398
Db 1321 ATCAAGTGAACAGCGATTTCTCACTCATTCGTGTGCTGCGAGGATAGGCAAGTAACTTG 1380
Qy 1399 GATATGGGCTTCCGCTTAGGCTCAAGCTGCTGCTTAAACCACTGTCAAAACCCGACCG 1458
Db 1381 GAGATCGGGCTGAGAACAAAGATTCAGAAAATGTTTAAAGGACGCAAGAGGCCCTCACT 1440
Qy 1459 GCTATTAACAATGCGCATGTGAGACATCTGCGTGGCTTTACACAGAAAGCTGAAGAAATG 1518
Db 1441 CTCAATACCCCGAGGAGCTGAAGAAAGTAAAGTCCAGCCGATGAGGCTTAAGAGAGTGT 1500
Qy 1519 GCTGCAAGGAAAGATTCAGAAAGCCCTGCAACCTTGTGCTCCCTGAATTAAGAAAAGAG 1578
Db 1501 CGTAAAGCCAGAGGAGTGTGCGGCGAGCTTACCACTTTTGGCAGCTGATGTGAAGAGCC 1560
Qy 1579 ACCGTAGAGCGAGAAATGACCTCATTTATGCAAGAGCGAGAGCGAGTACGCTGAGACA 1638
Db 1561 ACTGTGAAGCCGATGTGCACTTGTATGTAACAAAGGCTGGGCGCGCTCAGTGAAGACA 1620
Qy 1639 CCAGAGAGACATCAGGCTGAACAACTTACCAGGCGCAAGAAAGATTTGGCTTTACGCT 1698
Db 1621 CCTGCTGGCTTGAATTAAGTTAACAGCTACGCTGGCGAGGACAAAGATCGGCTTTACGCT 1680
Qy 1699 ATACTTTACCCCGAGGCTGATTAAGTATGTAAGAAATACTGGCGTATTCACCCCTTGGCG 1758
Db 1681 GTGCTTTCTCCGAGGCTGTATCTCAAGAGTAAATTAATCTTGCATCTCACCTCTGCT 1740
Qy 1759 GAACAACTACTGTGTAATGACTCAACAAAGTAAAGGAGGAGTAAAGTGCAGACCTATAC 1818
Db 1741 GAACAACTCAATGTAATTAACACACTCTGGCGGAAAGGGGTTATGCGGTGAACCATAC 1800
Qy 1819 CACGTAAGCTATTTGAACGAAGGACGCGGCTCCCTGTTCAAGCTTCAGGCAATTG 1878
Db 1801 CATGTAAGATGAGTGGCCAGAGGAGCATGCAATACCCGCTCCAGGACTTTCAAGCTCTG 1860
Qy 1879 AGTGAAGCGCTAGATCGTTTTCACAGAGAGGAGTTCGTAACAAATCTGTCACAC 1938
Db 1861 AGTAAAGTGCACATTTGTGTACAAAGACGTAGTTCGTAACAGATCTGACACAT 1920
Qy 1939 ATCGCAATCAACGAGAGGCGCTTAACACTGACGAAAGTACTTAAGACTGTAAAGACT 1998
Db 1921 ATTGCCACATGAGAGGAGGCTGAACACTGATGAAGAAATTAACAAACGTGCAAGCC 1980
Qy 1999 CAGGACACAGCTCAAAATAGCTTCTGATTAATGACGACGAAAGTGTGAAGCGAGAA 2058
Db 1981 AGCAGACGACGCGCGAATACCTGTACGACATCGACAGAAACAGTCCGTCAAGAAAGAA 2040
Qy 2059 GAGCGAGTCCCTTGTGCTTAACCGGTGATCTGTGATCAACCATTTCAAGAGTTGGC 2118
Db 2041 CTATGACTGGGCTTACGAGGCTTACAGGCGAGCTGTGATCTCTTTCATGAATTTGCC 2100
Qy 2119 TACGAGAGCTCAAGACACGACGACAGCACTTCAACAAAGTCCCAACCATGAGTAT 2178
Db 2101 TACGAGAGCTGAGAACACGACGAGCGCTCTTACAAAGTACCAACATAGGGGTAT 2160
Qy 2179 GGAAGTGCAGGTTCAAGTAAATCTGAAATCATCAAAAGCGCTGTACTAAGAAAGATCTG 2238
Db 2161 GGCCTGCAGGATCAGGCAAGTCTGGCATCTTAAAGCCAGTCAACAAAAGATCTTA 2220
Qy 2239 GTTGTAGTGTGGAAGAAAGAACTGGCGCAAAATCACTCAGGATGTAAGAGAGATGAGA 2298
Db 2221 GTGTGTAGCGCGCAAGAAAGAACTGTGCAAGAAATTAAGGAGCTCAAGAAAATAGAAA 2280
Qy 2299 CGTATGATGTGTGCTAGGACTGTGATTCAGTGTCTTCAATAGGGGTTAAGCAACCC 2358
Db 2281 GGGCTGACGTCAATGCAAGACGTGTGACTCACTGCTCTTGAATGATGCAACACACCC 2340
Qy 2359 GTTAACACTCTGTATCATTTGATGAGGCAATTTGCTGCTGATGAGGAGCGCTGTGCACTG 2418
Db 2341 GTAGAGACCCGTATATTGAAGAGCTTTGTGCTTGTATGATGAGGATCTTCAGAGCGCTC 2400

QY	2419	ATTGCGATGCGCAAACTTAAGAAAGTGATATTGCGCGGGACCCAAAACATATGGCGCTC	2418
Db	2401	ATAGCGATTTATPAGACTCTAAAAAGCACTGCTCTCGGGGGATCCCAACACATGGCGTTT	2460
QY	2479	TTTAAACATGATGCGCTTAAGAAAGTACATTTTAACCATATGACATATATGACATGAAGTACAT	2538
Db	2461	TTTAAACATGATGCGCTTAAGAAAGTACATTTTAACCATATGACATATATGACATGAAGTACAT	2520
QY	2539	AAAAGCATCTCTAGAGAGTGCAACACAGCTGTAAACCGCATGCTCTCAACGCTCTTCTAC	2598
Db	2521	AAAAGCATCTCTCGCGCTTGGACCTAAACTGTGACTTGGGTCGTCCTCAACCTTGTTTTAC	2580
QY	2599	GACAAAGCAATGGAAGACGGTTTAACCATATGCTGTATTAATATCATATGATATACACAGGG	2658
Db	2581	GACAAAAAAATGAGAACACACGATTCGAAAGAGACTTAAGATTTGATTTGATGACATCAACGGC	2640
QY	2659	ACCAACAAAGCCGACAAAGATGATCTGATTTCTAACCTGTTTCAGAGATGGGTGAACAG	2718
Db	2641	AGTACCAAACTTAAGCAGAGAGATCTCATTTCTCACTTGTTCAGAGGGTGGGTGAAGAG	2700
QY	2719	CTACGATTTGACTACAAAAATCAGAAATCATGACTGCGGCTGCATCCCAAGACTTACG	2778
Db	2701	TTGCAAAATGATTTCAAAAGCACAACAAATTAATGACGCAAGCTGCTCTCAAGGGCTGACC	2760
QY	2779	CGAAAGGCGTTTATGCTGACAGTACAAAGTCAACGAGATTCACATCTACTATCCGAGACT	2838
Db	2761	CGTAAAGGTGTGTATGCCGTTGCGTACAGGTAATGAATAATCCTGTATACGACCCACC	2820
QY	2839	TCTGACACGTGAACGTGTACTTACACGACAGAAAAACGATTTGTGMAACGCTA	2898
Db	2821	TCAGAACATGTGAACGTCTCTACTGACCGGACGGAGGACCCGATGTGTGAACAACTA	2880
QY	2899	GCTGTGATCCCTGATTAAGAAGACTTACAGTAATATTCGCGGGATTTACGGCTTCA	2958
Db	2881	GCCGCGGACCCATGATATAAAAACATGTGCTGCAAGTACCTGTGGAAATTTCTACGTCACG	2940
QY	2959	TTGGACGACTGGACGCGGAAACAGACGCCATTTAAGCAGCGTTTGTATPAGCCGAG	3018
Db	2941	ATAGAGAGTGGCAGACGAGCATATATCCATCATAGGCACTATTTGAGAGACCCGAC	3000
QY	3019	ACAGCTGATGTGTTCCAGAAATPAGGTGAACGTCTGCTGGGGAAAGGCTTTTGAAGCCAGTC	3078
Db	3001	CTTACCGGACTCTTCCAGAAATPAGGCAAAACGTTGTGGGCAAGGCTTTTATGTCGCGTG	3060
QY	3079	TTGGCAGCGGCCAACATTTGTGCTGACGAGACAGAGTGGGAGAGCTTGACCCATTCAG	3138
Db	3061	CTGAAGACCGGTGGATGACATGACCTGAAACATGGAACACTGTGATTAATTTTGA	3120
QY	3139	CATGACAGCGTACTCACTGAAATGSCACTGAACTTTTGGACACAGGTCCTTTGA	3198
Db	3121	ACGACAAAGCTCACTACGACAGATGATATTGAACCAATATGCGTAGGTTCTTTTGA	3180
QY	3199	GTAGACCTGGAACAGTGGGTATATTTCCGCTCTCAACGTCGCACTTACAGGGAATCAG	3258
Db	3181	CTCGATCTGGAACCTCGGTCTATATTTCTGACCAACTGTTCCGTATTCATTAAGGAATAT	3240
QY	3259	CACGTGGATTACTCGCCAGGGAAGAACATGTATGGGCTTAATAGAGAGTGCAGAAAGAG	3318
Db	3241	CACGTGGATTACTCCCGCTGCTTAACATGTACGGGCTGAATTAAGAAAGTGTCCGTACG	3300
QY	3319	TTGTACCGGCAATTCGCTGATCAACAAGCGGTTTGAACAGGCGAGGGTAGCTGATTA	3378
Db	3301	CTCTCTCGCAGTATCCCAACTGCTCTGGGCAATGTCGCACTGGAAGAGTCTATGACATG	3360
QY	3379	AGGAATTAATCCATCAAGAGACTCTCTCAACAATTAATGTGTTCCATTAATATGCCGG	3438
Db	3361	AACATGTGTACACTGCGCAATTAATATATTCGCGCATTAATCTGTAAACAGAGA	3420
QY	3439	TTGCCCACTGTTGATGTTGTGACCAACAAGACAGGGTACACATGATCAACAGCGATTC	3498
Db	3421	CTGCTCATGCTTATGTTCTCCACCATATATGAACCCACAGAGTGACTTTTCTTTCATTC	3480

QY	3499	CTATCTAAGAAAGGGGCAATCTGTGTGGTAGTCGGCATCTCTATCAGATTCGAGG	3558
Db	3481	GTCAACAATTGAAGGGAGACAGTCTCTGTGTGGGAAAGTTCTCGTCCAGGC	3540
QY	3559	AAGAAGTAGAGTCCATGGGTCCATTGGCCCACTATACATCAGGTGTGATCTCGATTTG	3618
Db	3541	AAATGTGTGACTGGTTGTCTAGACCCGGCTGAGGCTACCTTCAGAGCTCGGCTGATTTA	3600
QY	3619	GGAAATCCTAGCCATGTCCGTAATATGACATTATCTTTGTCAATGTTAGACCCCGTAC	3678
Db	3601	GGCATCCAGGTGATGTGCCAAATATGACATAATATTGTGTTAATGTAGAGACCCCATAT	3660
QY	3679	AGGAACCACTACTACCAACAGTGGAGATACGCTATCCACCAACAGACTAAACGT	3738
Db	3661	AAATACCATCTATACAGACTGTGAAGACCATCCATTAAGCTTTAGCATGTGTAAACAA	3720
QY	3739	AAGGTGTCCACCACTGAACATCTGGCGGAACATGTGTGGCTATAGGATAGGCTTGCT	3798
Db	3721	AAAGTGTGTGACTGTGAATCCCGGCGGAACCTGTGTCAAGTATAGTTATGTAGCCT	3780
QY	3799	GATCGCGCAACCGAATATCATCACTGCGGTGCACGCTCATTTAAGTTTACCCTGTGC	3858
Db	3781	GACAGGGCCAGCAAGACGATCATTTGTGTATAGCGCGGACGTTCAAGTTTTCGGGTA	3840
QY	3859	TGTCAAGCTTAAGAACATCGCCGAAATATCTAGGTTCTCTTCGTGTTCTTGGCAAGAC	3918
Db	3841	TGCAAACCGAATCTCATCTTGAAGAGACGGAAATTTCTGTTGTATTCATTTGGGTAGAT	3900
QY	3919	AACGGCAACCAACATATGACACAGACAGATCGGTGTAGTGCTGAACATGTATCA	3978
Db	3901	CGAAGGCCCTTAGCGACATCTTTACAGCTTTCATCATACCTTGACCTAACATTTATACA	3960
QY	3979	GGGTCAACCAAGTACGAGGCGAGGAGATCTCCAGCTTACAGAGTATCAGAGTGCATTT	4038
Db	3961	GGTTTCAGACTCCACGAAGCGGATGTGCACTCATATCATGTGTGTGGCAGGGGATATT	4020
QY	4039	AGCAAGACGCTGACCAAGCTATGTGTTATGTCTGTATATGCAAAAGTCAACCAAGTTCC	4098
Db	4021	GCCACGGCCACCGAAGGATGTATTAAATGCTCTTAAACAGCAAGAACCAACTGGCGGA	4080
QY	4099	GGAGTGTGCGGTGCACTGTATCCGAATATGCGCGCTGCTTTGTATPACAGCCAAATAGCT	4158
Db	4081	GGGTGTGCGGAGCGCTGTATPAGAAGTTCGCGAAAGCTTCGATTTACAGCCATGGA	4140
QY	4159	GTCGGGACCGGTACACTTGTGTGAAACAGAAC--GCTCATATACATAGCTGTAGAGACC	4215
Db	4141	GTAAGAAAGCGGCACTGTGTCAAAAGGTGTGACGTAAACATATCATTTCAATGCCGTAGACCA	4200
QY	4216	AATTTTCTAAGATGCGGAAACCGAGGCGACCTTAAGCTCGAGCTGCTCATAGAGC	4275
Db	4201	AACTTCAACAAGTTTGGAGGTGTAAGGTGACAAACAGTTGGCAGAGGCTTATGATGTCC	4260
QY	4276	ATACGCTTCATCTGTCAACGCTGAGCGGATTAACAATAATATCAACCGCTACTGTCAACC	4335
Db	4261	ATCCTTAAGATTGTCAACGATTAACAATTAACAAGTCAAGTCAAGATGAGATTCACTGTGTCAACC	4320
QY	4336	GGCATCTATTTGTGTGGCAAAAGATCGAGTGTGCAATCATTTGCGATCACTCTGTCACTGCT	4395
Db	4321	GGCATCTTTTCCGGAACAAAGATCGACTAACCAATCATATTGAACAATTGTCTGACAGCT	4380
QY	4396	TTGCAACACTACGATGCGGATGTACCAATATATTGCTTGTGATTAACAATGGGAGACGAG	4455
Db	4381	TTAGACACCACTGATGTGAGATGTAGCCATATATCTGAGGGGACAAGAATGGGAATGACT	4440
QY	4456	ATATTCAGGCTCATTCACCGCAAGAAAGCTCGAAATTTCTGGA-----T	4500
Db	4441	CTCAAGGAGAGAGTGGCTAGAGAGGAAGCAGTGGAGAGATATGACATATCCGACGACTCT	4500
QY	4501	GATGCAAGCAGTAGACATTGATCTTGTGTACGGGTTCACCCAAACAGCTCTTTGGCAGGC	4560
Db	4501	TCAGTGACAGACCTGATGACAGACTGTGTGGGTGTGATCCGAAAGATTTCTTGGCTGGA	4560
QY	4561	AGACCAGATTACTCCGTCAATGAGGGGCAAGTTGTATTCAATCCGTAAGGTACACGATTC	4620

Db 4561 AGGAAGGCGCTACAGCACAGCGATGCAAAACTTCTCATATTTTGGAAAGGACCAAGTTT 4620

Qy 4621 CATCAGACCGGCCAAGSAGCATTTGCCGAATTCATGCATGATGTGGCCCAACAATCTGAGGCT 4680

Db 4621 CACCAAGGGGCCAAGATATAGCAGAAATTTATGTCATGTGGCCCGTTGCACAGGAGCC 4680

Qy 4681 AATGAGCAGATTTGCTTGTATACATCCGTGGGGAGATATGTCAGCATCCGCTCCAAATGC 4740

Db 4681 AATGAGCAGATATGATATATATCTCGAGAAAGCATGACAGATTTAGGTCGAATGC 4740

Qy 4741 CCAAGTAGAGAGTCAAGGCGCTTGCTCCACTCACACATTCATGCTGTGTAATTAC 4800

Db 4741 CCGGTGCAAGAGTGGAAAGCCTCCACACACTGACAGCTGCTTGTGTCATCAT 4800

Qy 4801 GCTATGACGCGTGAAGCGCGTATATACAGTTGGCGCTTCGGAAGAAAGAACAGTTGCCGTA 4860

Db 4801 GCCATGATCTCAGAAAGATACAGGCGCTTAAAGCTCAGCTCCAGAACAAATTAATCTGTG 4860

Qy 4861 TGCATCATATTCCTGTGGCCGAGTACAGGATCACAGGCGTGGCAGAAAGCTACAGTGCAGC 4920

Db 4861 TGCTCATCTCTTCATTTGCCAAGATATGAATCACTGTGTGTCAGAAATCAATGCTCC 4920

Qy 4921 AAACAGTCCTGTTTTCAAGGGGTGTACACCGGCTGACACCCAGGAGATA 4973

Db 4921 CAGCCTATATTTGTTCTCACCGAAAGTGGCTGGTATATTCAATCCAGGAAAGTA 4973

CC	AAV74108 standard; cDNA, 11464 BP.
ID	AAV74108
XX	
AC	AAV74108;
XX	
DT	12-APR-1999 (first entry)
XX	
DE	Venezuelan equine encephalitis virus IE variant cDNA.
XX	
KM	VEE IE virus; vaccine; ds.
XX	
OS	Venezuelan equine encephalitis virus.
XX	
PN	W09853077-A1.
XX	
PD	26-NOV-1998.
XX	
PF	20-MAY-1998; 98MO-US10645.
XX	
PR	16-DEC-1997; 97US-0991840.
XX	
PR	20-MAY-1997; 97US-0047162.
DR	24-JUL-1997; 97US-0053652.
XX	
PA	(REED-) REED ARMY INST RES WALTER.
XX	
PJ	Crise BJ, Oberste MS, Parker MD, Schmurra SM, Smith JF;
XX	
DR	MPJ. 1999-045316/04..
XX	
PT	New DNA encoding infectious Western or Venezuelan equine encephalitis virus genome - useful for the production of live or attenuated vaccines for human or veterinary medicine
XX	
PS	Claim 24; Page 68-79; 112pp; English.
XX	
CC	This cDNA sequence codes for an infectious Venezuelan equine encephalitis virus IE variant (VEE IE) RNA transcript. Using primers (see AAV7412-23) specific to genomic RNA of VEE IE, RT-PCR was used to generate cDNA fragments which were subsequently cloned and used to assemble the full-length cDNA. The full-length infectious clone is useful in the production of virulent VEE IE virus, and for introducing and testing attenuating mutations. Also new are: infectious or attenuated VEE IE RNA transcripts and viral particles; cDNA (see AAV74107) encoding an infectious Western equine

CC encephalitis (VEE) viral genome; infectious or attenuated VEE RNA
 CC transcripts and viral particles; an attenuated chimeric virus (av)
 CC containing non-structural sequences from a first alpha-virus (av)
 CC and structural sequences from a second av, resulting in attenuation
 CC of the second av; a method for expressing a protein (especially an
 CC antigen to protect against a pathogen) by cloning its gene into an
 CC attenuated VEE or VEE IE replicon so that transcription of the
 CC replicon produces RNA able to infect the cells in which protein is
 CC to be produced; methods for diagnosing VEE and VEE infections; and
 CC polypeptides encoded by VEE variant IIA (see AA074109). Attenuated
 CC VEE and VEE are used in live or inactivated vaccines, for use in
 CC human or veterinary medicine. Chimeric viruses are also useful as
 CC vaccines, directed against the second av which is particularly
 CC Eastern equine encephalitis (EEE) virus or some variant of VEE.
 CC VEE and VEE nucleic acids are used as primers and probes to
 CC diagnose virus infections and to define natural variants, also for
 CC production of protein antigens which can be used as diagnostic
 CC reagents, to generate antibodies, and in vaccines. The attenuated
 CC viruses are highly immunogenic and provide long-lasting protection.

XX
 SO Sequence 11464 BP; 3259 A; 2734 C; 2900 G; 2571 T; 0 other;

Query Match Best Local Similarity 64.1%; Score 2074.8; DB 20; Length 11464;
 Matches 3202; Conservative 0; Mismatches 1772; Indels 24; Gaps 4;

QY	1	ACCTTACAAACTAATCGATCCAAATTTGGAAAGAAATTCAGGTTGACTTAGACTGTACAGC	60
Db	20	AGCCCAAAACCATATACTACCCAAATATGGAGAAAGTTACAGTTGACATTCAGAGAAAGATAGT	79
QY	61	CCGATATGCAAGTCCTTACAGGGACGTTTCCACAATTTGAGATGGAAGCAAGCGAGTTC	120
Db	80	CCCTTCCCTCAGAGCAATTACAAGGAGCTTCCGCAgTTTGAGAGTGAAGCCAAAGCAGGTC	139
QY	121	ACTGCATATGACCATGCGCAATGCCAGAGGTTTTGGCATGTGSCAACAAAAGCTCATTTAG	180
Db	140	ACAGATTAATGACATGCTTACGCCAGAGGTTTTGGCATTTGSCATTCGAAATTTGATCGAG	199
QY	181	AGCGAAGTCGACCGGGACCAAGTTATCTTGGACATTTGGAAGTGCGCCCGCTCAGACATGCA	240
Db	200	ACGGAGGTGGAAACCATCCGATACATTCCTAGACATTTGGAAAGTGGCGCTCCTCCGCAAGATG	259
QY	241	CATTGCAATCACCGCTATCATTTGTATCTGCCCCATATGATTAAGCGCTGGAAGACCCCGACAGA	300
Db	260	TATTCCAAGCACTAAGTACCATTTGACTCTGTCCGATGAAATGCGAAGAAATCCCGACAGA	319
QY	301	CTACAACGSTATGCGAAGAAAGACTTAAAGAAAGT-----GACATTACCGACAAAGAACATA	354
Db	320	CTGTTTAAGTATGCAAGCCAGCTGAAGAGAACTGTAAAGAGATTACAGATTAAGGAACCTG	379
QY	355	GCCCTTAAAGCGCGCAGACCTGCTGGAAgTCATGTCAACAACGACGACAGACTCCATCT	414
Db	380	GACAAAGAAAGATGAAGAGAGCTTGGGAAGTCATGAGGAGACCTGTACTCGAAACTGAAACG	439
QY	415	CTGTGTATGCAACAGACGCCACGTGTAGGTACTTTGGAAGGTAGCACTATATCCAAAGT	474
Db	440	ATTTCCTTTCACGACGATGAAACCTGTGATTTGAGGGCTCAAGTCGCAGTGTATCGAGAT	499
QY	475	GTATACGAGATCATATGCAACCGACATCAATCTACCAACAGGCGCTTAAAGAGTTAGAGCA	534
Db	500	GTTATGCGGGTTGACGACCGACCGAGGCTTTTCCATCAGGCCCAACAAAGGGCTCAGAGTC	559
QY	535	ATTTACTGATAGGCTTTGACACGACCCCTTTTATGTACAAAAACATGSCAGGTTCTTAC	594
Db	560	GCGATTGGAATAGATTCGACATACCCCTTTTATGTTTAAGAACCTGTGCTGAGCATAT	619
QY	595	CCTACTTACAAACGAACTGGGCTTACCGAGAGATTTGGAAGCAGTAACTATGGCCCTC	654
Db	620	CCCTCTTATTCGACCAACTGGGCCACGACGACCGGTTTAACGGCTCTGTATATATAGCCTTG	679
QY	655	GGTACTCAGATCTTCAGAGAGCAGGGCTTGGAAAACTCTCAATCCTTTGGAGAAAGAGAG	714
Db	680	TGCACCTCCAGATGATGAGCGGTTCACGTAAAGGAGATCTCAATCTTCAGAAAGAAATTT	739

QY 715 CTCCACTACTAATAGATCATATTCGTGGTTCACAACTATACAGAGATGA 774
DB 740 TTAATACCATCAATATAGTCTTGTCTCTGTAGATCTACATCTACACAGAGAGCA 799
QY 775 TCACGTTCAGTAGCTGGCATTTCCAAAGTGTTCACCTGAAAGAAAGTCTAATTC 834
DB 800 GACTTACTAAGAGATGGGACCTACCGTCCGTTTTCACCTACGTGTGAAGAGAAATTAC 859
QY 835 ACAGTATAGTGGGACCATTTCTCAGCTGTGAAGGTATGTATCAAAAAGTAAAGATC 894
DB 860 ACTTCTGGTGTGAACATAATAGTGTGCGAGCGGTACGTGTCAAAAAGTAACTATT 919
QY 895 AGCCAGAGACTATACGTAAAGTTGAGAACTTGGCGTCCAAATGATCGAGGGGTTTC 954
DB 920 AGTCCAGGTCTGTACGGGAAACCGTCCGGCTACGCTGCCACATGCTACGAGGATTC 979
QY 955 TTGAGTTGCAAGTACAGATACGCTGCGGCGAGAGGGTTCTTTTGTGTGTACG 1014
DB 980 TTGTCTCAGAGTACGACGACGCTTGACGGGAGAGGGTCTCTTTCCCGTATGACG 1039
QY 1015 TATGATACGACCACTTTGGATGAGATGACAGGATTTCTGGCACTGACGTATGTG 1074
DB 1040 TACGTCAGCCACATGTGTGACCAAAATGACAGGCAATTCG3CAACAGATGTACGTGA 1099
QY 1075 GATGACGACAAAACATAATTTGGTGGCTCAACCAAGATTTGTGTCAATGTAGAGC 1134
DB 1100 GACGACGGGCAAAATTTGTGTGGTCTCAACGCAATGTGTAAATGTGGCCGCACT 1159
QY 1135 CAAGAATACTATACACATGACAGATCTATTTACAGTGTGTGCGCCAGCGTTTTC 1194
DB 1160 CAAGAATACTATACACATGAGAACTATTTATACCGTGTGTGACCAAGCATTTGCC 1219
QY 1195 AGTGGGGCGGGAATCTGTGCGCATTTGAGACGACGAGAAAGAACTAGGGTGGCGAG 1254
DB 1220 AGATGGGTGAAGATACACAAAGATCAAGAAAGCAATTTGGGGCTTAGGAGAC 1279
QY 1255 CGCACTCTTACTATGGGCTGTGGCTTTCAAGACCCAGAAATCATCATCTAC 1314
DB 1280 CGCCAGTGTGTAATGGGGTGTGTGGCGTTCAAGAAACACAAATACATCACTGATC 1339
QY 1315 AAGAACTGTGACGCAAAACATTAAGAAAGTACCTGCGTCTTGAATCTTGTGAT 1374
DB 1340 AAACGACGACACCCAAACGATCATCAAGGTAAACAGGATTTCACTTTGCTGTG 1399
QY 1375 CCAAGCTTACAGCAGCGGCTCATATGGGCTTCCGCGTGAAGCTCAAGTGTCTCT 1434
DB 1400 CCCAATTTGAGAGCAACCTTTGAGATTTGGCTGAGGACAGAGATCAGAAATCTACTG 1459
QY 1435 GAACCACTGTCAACCCGCACTGATTAATGCGAGTGTGAGATCTGGGTGC 1494
DB 1460 GAGAACTGTGACAGACACCACTGATTAACCGCGACGATACAGAGAACCAAGAC 1519
QY 1495 TTACGACAAAGAGTGAAGAGTGTGCTGACGCGAAGAGATCAGAGAACCTGCGACCC 1554
DB 1520 GCGGGGATGAGGCTTAAGGAAGTCAAGAGGCGGAAGGCTCAGGGCAGCATTTACACCG 1579
QY 1555 TTGCTCCTGAATGAAAAAGAGACCGTAGAGGAGAAATGACCTCATTAATGCAAG 1614
DB 1580 CTGTCTGCGATGTAGAGAACTGCTGAGGCGGAGGTTGACTTAATGCTGAGAG 1639
QY 1615 GCAGAGAGAGTGTGAGAGACACAGAGAGCATCAGGGGTACAAAGTTAACCCAGGC 1674
DB 1640 GCGGAGAGAGATCTGTGAAACCCCGAGGGGCTCATCAAAATCAACGATTAACAGAG 1699
QY 1675 GAAGAGAAATTTGGTCTTACGCTATCTTCAACCCAGCGGCTTTGAATAGTAAAA 1734
DB 1700 GAAGCAAAATTTGCTTTATGCGGTGTCTCCCAAGCAGATCTGGAAGTAAAA 1759
QY 1735 CTGGGTGTATCCACCTTGGCGGAGCAAGTATGTGATATCACTCAAAAGTATGCGCA 1794
DB 1760 CTGAGGTCACTCAACCCGCTTGACAGAGCAAGTAATTTATATCACTCTGGAAGAG 1819

QY 1795 GGGAGATACAAAGTGCAGCATATACACCGTAAAGTCAATGTATACAGAGAGCGCGTCC 1854
DB 1820 GGTATATGCGAGTGTGACCTTTACACGGAAGAAAGTGTAGTATCCGGAAGGCAACTATA 1879
QY 1855 CTTGTTCAAGACTTCCAGGCAATGATGAGAGCGCTACGATCGTTTCAACGAGAGGAG 1914
DB 1880 CTTGTCAAAGACTTCCAGGCACTCAGTGAAGCGCCCAAAATGTGTACAAAGAGTGTAG 1939
QY 1915 TTTCGTAACAGATACCTGCAACATTCGCATCAACGAGAGAGCCCTAAACCTACGAA 1974
DB 1940 TTTCGTAATGATATTAACACCAATTCACGACGCGGTGAGCTCTGAACACAGCGAG 1999
QY 1975 GAGTACTATAAGACTGTAAAGACTCAGGACACAGACTCAGAAATAGCTCTGATATGAC 2034
DB 2000 GAGTATTAACAGAGTGTAAACCTAGTAAATAGAAAGTGTACTTATATGACATTGAC 2059
QY 2035 GCAAGAAAGTGTATTAAGCGAAGACGACAGTCCCTGTGCTTAACCGGTATCTGTA 2094
DB 2060 AAGAAACATGTGTGAAGAAAGAACTGTTACGGGACTAGGTTTGAACAGAGAACTTGT 2119
QY 2095 GATCCACATTTCAACGATTTGCGTACGAGAGTCTCAAGACACGACGACACCTCAC 2154
DB 2120 GATCCCTTCCATGAATTTGCAATATGAAGCTTGAAGACAGTCCGCGCAGCTTAT 2179
QY 2155 AAAGTCCCAACATCGAGTCTATGAGTGCAGAGTTACAGTAAATCTGAATCATCAA 2214
DB 2180 CAAGTACCACTATATAGTGTATAGCTTCCGGGTCCGGAGTCTGGAATTTATAG 2239
QY 2215 AGCGTGTACTAAGAAAGATCTGGTTGTGAGTGGAGAAAGAACTGCGCAGAAATC 2274
DB 2240 AGCGAGTCAAAAGAAAGACTTGGTAGTGTGCTAAGAAAGAAAGAACTGCGCGAGATA 2299
QY 2275 ATCAGGATGTAAAGAGATGAGAGTATGATGTGCTGCTAGAGACTGTGATTCAGT 2334
DB 2300 ATAAAGGAGCTCAAGAAATGAAGGCTTAACGTCATATGCCGAGCGGTGACCTCAATG 2359
QY 2335 CTTCTAAATGGGGTTAAGCAACCCGCTTAACCTCTGATATGAGAGCAATTCCTGC 2394
DB 2360 CTGTTAAATGATGACAGACACCCGCTCGAGACTTTGTATGATGAGACCTTTGCAATC 2419
QY 2395 CATGAGGAGCGTGTGCACTGATTTGCCATTCGTCAAACTGAAGAAAGTGTATGTGC 2454
DB 2420 CACGCGGCACTCTCAGGGCTTAATAGCAATTAATAGCCCAAGAAAGAGTGTATGT 2479
QY 2455 GGGACCCAAACAAATGGGCTTCTTAAACATGATGTGCTGAAAGTCAATTTAACAT 2514
DB 2480 GGTGACCCAAAGCAATGTGGCTTCTTCAACATGATGTGCTGAAAGTGCACCTTCAACAT 2539
QY 2515 GACATATGACATGAAGTGTACATTAAGAAAGCATCTCTAGAGAGTGCACAGACTGTAC 2574
DB 2540 GAATATGACATCAAGTTTTCACAAAGCAATATCTCGAGGTGTACCAAGTCAAGTACG 2599
QY 2575 GCAATCGTCTCAGGCTCTTACAGCAAGAGAAAGAAAGGTTAACCCATGTGTAT 2634
DB 2600 TCGGATGTGTCACAGCTGTTTATGACAAAGAAATGAGAACTATCAATCAACCAAGGATTC 2659
QY 2635 AAAATCATATGATATCAACAGGACCAAAAGCCGCAAAAGATGATGTGATTTAAC 2694
DB 2660 AAAATCGAATTTGACAAACAGAAAGTACTTAACCAAAAAAGATGACTTATCTCA 2719
QY 2695 TGTTCAGAGATGGGTAACAGCTACAGATTGCTCAAAAAATCAGAAATCATGACT 2754
DB 2720 TGTTCAGAGATGGGTTAAGCACTGCAATTAAGCTCAAAAGAAATGAATTAATGACC 2779
QY 2755 GCGGCTGATGCAAGACTTACGCGAAGAGCGTTATGTCTCAGGTCAAGTCAAC 2814
DB 2780 GCGGCTGCTCAGAGGATGACGCGAAGAGTGTCTATGCAATGAGTAAACAAAGTTAAC 2839
QY 2815 GAGAACTCATCTTACTGCAAGCTTTCTGAGCAGCTGAACGTTTACTTACGACGAGAA 2874
DB 2840 GAGAAACCATATGTATGCAACCACTCAGAGCATGTAATATGTCTGTGACCCGAGCGAA 2899
QY 2875 AAAGCATGTGTGGAAGAGCTAGCTGTGTATCCCTGGAATAAAGACCTTAACGCTAAA 2934

```
Db 2900 GACAAAGATTGTGTAGAGACTCTTGACAGGGGACCCTGTGATTAAGACCTTGACCGCGAAG 2959
Qy 2935 TATCCCGGGGATTTACAGGCTTTCATTGACACATCGGACGGCGAACAACAGCCCATATATG 2994
Db 2960 TACCCCGGAGATTTCTACATGCAACATGGMAATGGCGAGGCAACATGATGCATCATATG 3019
Qy 2995 GCACGCGTTTGTGATTAAGCCGACAGCATGATGTTCCAGAAATTAAGGTCAACGTCGCG 3054
Db 3020 AGACACATCTGTGGAGAAACCGGATCCACGAGATGTCTTCCAAAATTAAGCTAATGTTTGC 3079
Qy 3055 TGGGCGAAGGCTTTAGAGCCAGTCTTGCCACCGGCCAACATTTGCTGACAGACAGACAG 3114
Db 3080 TGGGCAAGGCACTTGATCTGTGCTTAAGACAGCGGGGATAGATTGACACAGAGACAG 3139
Qy 3115 TGGGAGACGTTGACCCATTTAAGCATACAGAGGCTACTCACTGAAATGCACTGAAAC 3174
Db 3140 TGGAAACACAGTGAATTACTTCAAGAGATTAAGGCCCTCAGCTGAGATTTGCTGAAAT 3199
Qy 3175 TTCTTTGACACAGAGTTCTTTGAGTAGACCTGACAGTGGGTTATTTTCCGCTCTAAC 3234
Db 3200 CAGCTGTGCGTGCATTTCTTGCTTAGACTTAAGTTTGTGTTTTCGCGCCCAACA 3259
Qy 3235 GTGCACTTACTTACAGGGATCAGACCTGGGATTACTGCGCGGAAAGAACATGTATGGG 3294
Db 3260 GTTCACCTCTCATTTAGGAACAACATTGGGACAACCTCACCGTCACCCAACATGATGCGG 3319
Qy 3295 CTTAATAGAGAGTACAGCAAGAGATTGTACAGGGGATATCTGCTGATACAAAGCGGTT 3354
Db 3320 TTGAATCAGAGAGTGTGTCAGCAACTATACAGAGATACCTCAATTAACAGTGGCGTGG 3379
Qy 3355 GACACAGGACAGGTAGCTGATATATAGGAATATACATCAAGAGACTACTCTCAACAATT 3414
Db 3380 ACTACTGGAGAGCATACGACATGAACACCGGTACTTTGGCCAAATTATGATCCCGCATTA 3439
Qy 3415 AATGTGTTCCATTAAATCGCCGGTTGCGCCACTGTTGATCGTTGACCAACAAAGACAG 3474
Db 3440 AATTAGTACCGGTGAACCGTGTCTACCATGCTCTGTGACGCAACATGCTGATCAT 3499
Qy 3475 GGTACAACTGATCAGAGGATTCCTATCTTAAGATGAAGGGCAATCTGTGTGTGATC 3534
Db 3500 CCTCCAGTGAATTTTTCGCGCTTTGTCTAGTAAGCTTAAGGCAACACGCTCTGTGATTT 3559
Qy 3535 GGCATCCTATCAGCATTTCCAGGGAAGAAAGTAGAGTCCATGGGCTCCCACTAAT 3594
Db 3560 GGTAGAAGATGATTAATTTACGTAAAGACGCTAAGCTGTTATCTGAACACCTGATTTCT 3619
Qy 3595 ACCATCAGGTGTGATCTCGATTTGGGAATACCTAGCCATGTGCGTAATATGACATTATC 3654
Db 3620 ACTTTTAGGGGCGCCTAGATCTAGGCATACCCAATGAATACCGAAGTACGATATGCTC 3679
Qy 3655 TTTTGTCAATGTTAGACCCCGGTACAGAAACATACATACCAACAGTGGAGATACAGCT 3714
Db 3680 TTCTTAATGTMAAACAACAGTACCGCTACCACTACAGCAGTGTGAGAGAACACAGCC 3739
Qy 3715 ATCCACCAACAGCATGCTAAGCTGTAAAGGCTGTCCACCACTGAAACACTGCGGGAACATGT 3774
Db 3740 ATTAAAGTTGAGCATGTTGACCAAGAGGCTGCTGCACTGAAACCCCGAGAACTGCT 3799
Qy 3775 GTGCTATAGGATAGGCTTGTGATGCGCAACCGAATATCATCACTGCGGTGCA 3834
Db 3800 GTGAGCATTTGTTACGGCTATGCGGACCGGCGAGTAAACATCATATGTCAGTTGCT 3859
Qy 3835 CGCTCATTTAGGTTTACCCGTGTCTGTACGCTTAAGAACTGCGGAAATATCTAGAGTT 3894
Db 3860 CGGAGATTCAAGTTCTCGAGGGTATGCAACCGAAGGTGTCTAAGAGAGAACCGAAGTG 3919
Qy 3895 CTCTTCGTTCTTCCGCAAGGACAAACGCGCAACCAACATGACAGAGAGAGTCTGCGGT 3954
Db 3920 CTATTTGCTTCAATGGGTTGATGCTAAACGGAACCAATACCATACAGTCTGCC 3979
Qy 3955 GTAGTCTTGCAACATCTATCAAGGTCACAGGTACGAGGAGGAGAGTCTCAAGCG 4014
Db 3980 TCCACCTGACCAATATTTACACCGGCTCGAGGCTCCATGAAAGTGTGCTGCACTTGC 4039
Qy 4015 TACAGATGATCAGAGGTGACATTAGCAAGAGGCTGACCAAGCTATCGTTAATGCTGT 4074
Db 4040 TATCATATAGTGGCGGGGATATAGCTACATGCCACGGAAGAGTAAATGTTAATGCTGCC 4099
Qy 4075 AATGCAAGAGTCAACACGAGTTCCGAGTGTGCGGTGACCTGTATCCGMAAATGGCCGCT 4134
Db 4100 AACGCAAGGGCCACGCCAGGACGTGAGTGTGGGAGCTGTATACCGGAAGTACCCGAA 4159
Qy 4135 GCTTTGATAGACCCCAATAGCTGTGTGGACGCTGACTGTTGAA---GCAGAACCG 4191
Db 4160 AGCTTCATTTTACCAACCAATAGAAAGTGGGAAAGCTGATTTGTTCAAAGGTAATCTCAAA 4219
Qy 4192 CTATCATATGCTGTAGAGACCAATTTTCTTAAGATGCGGGAACCGAGGGGACCTT 4251
Db 4220 CATCTCATATGAGTGGGGCCGAATTTTAACAAAGTGTGAAAGTGAAGGTGACAAA 4279
Qy 4252 AAGCTCGACGCTGCTCATGAGCATAGCGTCCATCTCAACGCTGACGCGGATTTACAAA 4311
Db 4280 CAGCTGCAAGAGGTATGAACTATAGCCAGGATTTATATAGCAACAAATTATAGATCT 4339
Qy 4312 ATATCAGTACCGCTACTGTCAACCGCATCTATTCTGTGGCAAAATTCAGATGATCA 4371
Db 4340 GTGGCTATTCGGCTTCTGTCCACTGGAATTTTGGCCGAAACAAAGATTAATGCAA 4399
Qy 4372 TCATTCATACCTGTTCACTGCTTGTGACACTACGGAATGCGCATATATATATGTC 4431
Db 4400 TCTTAACCATCTGTTAAAGGATGGAATGGAACAACAACGAGTGTGGCCATATCTGC 4459
Qy 4432 TTGATTAACAAATGAGAGACCAAGATATAGAGCCATTCACCGCAAAAGAACGCTGA- 4490
Db 4460 AGAGCAAGAAATGAGAGAGTGAAGCTTGAAGAGGTGTAGCAGAGAGAGGCGGTAGAG 4519
Qy 4491 -----AATTGTGATGATGACACAGCCAGTACATTTAATCT-----TGCTAGGCTC 4536
Db 4520 GAGATATGATCTCCGAAGATCTCTCCGTAGCAGAGCCGATGAGAGCTGTGTAGATT 4579
Qy 4537 CACCCAAAGCTCTTTGGAGGAGACCGAGTTACTCCGTCAATGAGGGCAAGTTGAT 4596
Db 4580 CACCTTAAGGCTCTTTGGCTGGAAGAAAGTTACAGCACTAGCGATGGGAAACATTC 4639
Qy 4597 TCATACCTGGAAGGTATACCATTCATCAGACCGCCAGAGCAATTCGGAATTCATGCA 4656
Db 4640 TCATATCTTGAAGGAACCAATTTTCATCAGGCGCGGAAGACATGCAAAATTAACGCT 4699
Qy 4657 ATGTGCCCAACCAATTTGAGGCTAATGAGCAATTTGCTTGTATCATCTGCGGAGAGT 4716
Db 4700 ATGTGGCTGCGCTACAGAGGCTAATGAGCAGGTGTCTTATATCATTTGCGGTGAAGT 4759
Qy 4717 ATGTCCAGCATCGCTCCAAATGCGCAGTAGAGAGTCAAGGGGTGTGCTCCACCTCAC 4776
Db 4760 ATGAGCAGTAAAGATCCAATGCGCCGTTGAGAGGTCAGAGGATCACCCCAAGT 4819
Qy 4777 ACACTTCATGCGCTGTGTAATTAACGCTATGACGCGCTAGCGGCTATACAGTTGCGCTCT 4836
Db 4820 ACATGCTCTTGTGTGTGATCACAAGCTATGACCCCGGAAGGGTTACGCGTTGAAGCC 4879
Qy 4837 GCGAAGAAAGACATTTGCGCGTATGCTCATCTCTGTTGCCGAAGTACAGATAC 4896
Db 4880 TCCCGCCCCGAAATTAACAGTTGTTCTTCTTCCATTTGCGCAAGTACAGATTAACA 4939
Qy 4897 GCGGTGAGAGCTACAGTGCAGAAACAGTCTCTGTTTCAGGCGTGTACCAACCGGCT 4956
Db 4940 GGAAGTGAAGAGTTCATATTTGCAATCTTATCTTCTCTTAAGTACCTGAGTAC 4999
Qy 4957 GTACACCCCAAGAGTAC 4974
Db 5000 ATACACCTTAAGAAATAC 5017
RESULT 8
AAx78130
```

AA78130 standard; DNA; 11060 BP.
AA78130;
20-AUG-1999 (first entry)
psfV link DNA.
Alphavirus; replicon; expression vector; heterologous splice site;
RNA splicing; viral RNA replication; antigen; vaccination; therapy; ss.
Semliki forest virus.
Synthetic.
MO9925859-A1.
27-MAY-1999.
13-NOV-1998; 98MO-CA01065.
14-NOV-1997; 97US-0065793.
(CONN-) CONNAUGHT LAB LTD.
Parrington M;
WPI; 1999-385140/32.
New vector (EV) used to express an antigen for vaccination, e.g.
against human immune deficiency virus
Disclosure; Fig 4A-H; 60bp; English.
This invention describes the construction of a novel expression vector
(EV) containing an alphavirus replicon having a heterologous splice site.
The expression vector contains (i) a DNA molecule (I) complementary to
at least part of an alphavirus RNA genome, and (ii) inserted into a
region of (I) that is not essential for its replication, a heterologous
DNA (II), under control of a promoter. (I) includes at least one
heterologous splice site (HSS) to prevent aberrant RNA splicing of the
alphavirus and is the complement of the complete alphavirus genomic
region essential for replication of viral RNA. The expression vector is
used to express (II) in humans or animals, e.g. to express an antigen
for vaccination (against human immune deficiency virus) or to produce a
therapeutically active protein or peptide. Introducing an HSS: (1) makes
it more likely that any splicing will occur at this site, rather than at
a cryptic splice site; (2) restores function of the alphavirus when
eliminated, and (3) may improve transport of RNA from the nucleus.
Sequence 11060 BP; 2978 A; 2826 C; 2926 G; 2330 T; 0 other;
Query Match 31.3%; Score 1565.6; DB 20; Length 11060;
Best Local Similarity 58.3%; Pred. No. 0;
Matches 2905; Conservative 0; Mismatches 2039; Indels 36; Gaps 8;

336 CCTATGCCAGCGCAGAAAGACCCCGAAGGCTCGATAGCTACGCAAGAACTGCGACGG 395
325 AAGAAAGTGAATTAACCGACAAAGAACTAGCCTTAAGGGCGGAGACTGCTGGAAGTC 384
396 GCTCCGGAAAGGTCGTGATAGAGATCGAGAAATATCAGCACTGAGACCGTC 455
385 ATGTCAACACGACGACGAGAGACTCTCTGTGTATGACACAGACGCGCATGTAG 444
466 ATGCTACGCGCAGAGAGCTGAATCTTCAACCTTTGCTGTGATACAGAGGTACGTGTGT 515
445 TACTTGAAGTGTACGATATACCAAGATGTGTACGAGTCAATGCAATCAATC 504
516 ACGCAGCGCAAGTGGCCGTATACCAAGACGTGTATGCTGTACATGCAACCAATCGCTG 575
505 TACCAACGAGCGCTTAAGAGGTTAGCAATTTACTGATAGGCTTTGACACGACCCCT 564
576 TACCATCAGGCGATTAAGGTGTCAAGAGCGCGTATGATGGTTGATACACACCCG 635
565 TTTATGTACAAAACATGAGGAGTTCTTACCTACTTACACACGAACTGGGCTGACGAG 624
636 TTTATGTGTAGCGGCTAGCAGGCGCGTATCCAACTACGCAACAACTGGGCGACGAG 695
625 AGAGTATTGAAACAGTAACTTGGCTCGGTAACTAGATCTTCAAGAGACAGGCTT 684
696 CAGGTGTACAGGCGCAGAAACATAGGACTGTGTGACGATCCTTGACTGAGGGAAGACTC 755
685 GGAAACCTCTCAATCCTTAGAAGAGAGGCTCCAACTACTAATAATATATATTCG 744
756 GCAAACTGTCAATCTCCGCAAGAAAGCAATGAAACCTTGCGAACAGTCACTGTTCTCG 815
745 GTTGTTCAACATCTACACAGAAAGATATCACTGTATAGTATGCTGCAATCTTCAAC 804
816 GTAGATATCACTGTATACGTAGAGCAAGAAAGTACTAGAGAGCTGCACTTACCTCC 875
805 GTTTCCTCACTTGAAGAAAGTCTAATCTTCAAGGTAGATGTGGACCAATGTGCACTGT 864
876 GTATTCCACTGAAAGGTAAACAATCTTTACTGTAGGTGATACCATCTGATCAATGT 935
865 GAAGGTGTCATCAATAAAGATTAACGATCAAGCCGAGACTATAGGTAAAGTTAGAAC 924
936 GAAGGTACGATTAAGAAATCAATGTGCCCCGCTGTAGGTAAACGTTAGGG 995
925 TTGGCTCACAAATGCATCGGAGGTTCTTGAATGTCAAAGTACAGATACGTGCGC 984
966 TACGCGTACGATACACGCGAGGATTTCTAGTGTCAAGACACAGACACTGTCAA 1055
965 GCGGAGAGGTTCTTTTGTGTGTGTACGTATGTACAGCCACACTTGTGATCAGATG 1044
1056 GGAGAAAGAGTCTCAATCCCTGTATGACACTAACGTCCTCAACCATCTGTGATCAATG 1115
1045 ACAGGATTTGCGCACTGACGTATGTGATGACGACAAACCTATGTGGGCTC 1104
1116 ACTGCAATCTAGACGACCGAGCTCACACGAGAGCGACAGAAATTTGTATGTGGAATG 1175
1105 AACCAAGATTTGTGTCATGTGTAGAGCGCAAAAGAAATCTTACACATGCGAATAT 1164
1176 AATCAGAGATATGTGTGAAGAAACAGACGAAACCTTACACATGATGAAGACTAT 1235
1165 CTATTACAGATGTGCGCAGGCGTTTTCAGGTGGGCGCGTGAACATGTGCGACCTTG 1224
1236 CTGCTTCGATTTGTGGCGGTGCAATTTAGCAAGTGGGGAAGGAATTAAGGCAACCTT 1295
1225 GACGACGAAAGAACTAGGGGTGCGGAGCGCACTTTATCTATGGGTGCTGCTGGCT 1284
1296 GATGATGAAAAACCTGTGGGTGTCCGAGAGAGGTACTTACTGTGCTGTGTTGGCA 1355
1285 TTCAAGACCCAGAAATACATCTACAAAGAAAGCTGTGTAACGCAAAATTAAGAA 1344
1356 TTTAAAGAGAGAAATGACACATGTACAAAGAAACAGACACCAACATATAGTAGAG 1415
1345 GTACCTGCGCTTTTACTACTATTTGTGATTCACGCTTACAGGACGAGGCTGATATG 1404
1416 GTGCTTCAAGATTAATCTGTTGTGTCATCCGAGCTATGTGTACAGGCTCGCAATC 1475

1405 GGCTTCGCGTAGGCTCAGAGCTGCTGTTGAACCACTGTCAAAACCCGACCGGCTTT 1464
1476 CCACTCAGATACCGCATTTAAATGCTTTTGGCCAAAGAACCAACG---GAGATTAA 1532
1465 ACAATGGCCGATGTGAGCATCTGCGGCTTTACAGCAAGAAAGTGAAGAGTGC 1524
1533 CCTGTTCTGACGGGTGCTGACGCCAGGGATGTGAACAAGAGGAGAGGTTGAG 1592
1525 GCGGAAGATCAGAGAAAGCCCTGACCTTGTCTCTGTAATAATGAAAAAGACCT 1584
1593 GCCAGCTGACTAAGGAAGCCCTTACACCCCTGCTCCCATGCGCCGCGGAGACGG 1652
1585 GAGGAGAAAGTGAACCTATTATGCAAG-----GAGAGAGAGTGAAGGTGAGACA 1638
1653 GTCTGACGCTGACGTTGAAGAACTAGAGTATACGAGAGTGCAGGGGTGTGGAACA 1712
1639 CCAAGAGACACATCAGGGGTGACAAATTACCCAGCGAAGAGAAAGATTGGTCTTAC 1698
1713 CCTCGACGCGGTTGAAGTCAACCGCACAGCCGAACGATCTAAGTGAATAAGTGA 1772
1699 ATACTTTCACTCCAGCGGCTATTGAATGTGAAAACTGGCGTGTATCACCATTGG 1758
1773 GTTGTGTCCTCCGACGACCGTGTCTCAAGAGCTTCAAGTTGGCCCGCTGACCTT 1832
1759 GAACAGTACTGTGAATGACTCACTCAAAAGGTAGGGGAGGAGATCAAAAGTGCAG 1818
1833 GAGAGAGTGAATAATTAACATTAACGAGAGGCGCGGCTTACAGGTGACGAGAT 1892
1819 CAGGTAAGTCTATTGTAACGAAGGACGCGCGTCTCTGTTCAAGACTTCAGGCA 1878
1893 GACGGCAGGGGTCTACTACATGTGATCGGCANTTCGGTCCGTGAGTTTCAAGCT 1952
1879 AGTGAAGCGGTAGATGCTTTTCAAGAGAGGAGTTCGTAACAGTACTTGACATC 1938
1953 AGCAGAGCGGCTACTATGTGTACAAAGAGGAGTTCGTAACAGGAACTTATAC 2012
1939 ATCCGAATCAACGGAGGAGCGCTAAACACTGACGAGAGTACTTAAAGCTGTAA 1998
2013 ATTGCGCTTACGAGCGCTGCTGTAACACGAGAGGAACTACGAGAAAGTCAAG 2072
1999 CAGAGACAGACTAGATACGTCTTGTGATTTGACGACGAAGTGTGTTAGCGAGA 2058
2073 GAAGAAGTGAACGCGAGTACGTGTGACGATTAATAAATGCTGTGCAAGAGAG 2132
2059 GACGAGGTCCCTTGTGCTTAAACGGGTATCTGTGATTCACCATTTCAAGTTTGC 2118
2133 GAACCGTCCGGT 2192
2119 TACGAGAGTCTCAGACACGACGACGACCTCAAAAGTCCCAACCATCGAGTCTAT 2178
2193 TACGAGAGGCTGAAGATAGGCGCTGCGACCACTAATTAAGTACTAAGTATGAG 2252
2179 GAGGTGCCAGGTTCAGGTAATCTGGAATCATCAAAAGCGCTGTGACTAAGAAAG 2238
2253 GGGGTTCCGGGATCAGGCAAGTCTGTATTAATAAGCCTGTGACCAAAACGATCTG 2312
2239 GTTGTGAGTGCAGAGAGAAAGAACTGCGCAGAAATCATCAGAGATGAAGAGATGA 2298
2313 GTCAACAGCGGCAAGAGAGAACTGCGCAGAAATAGTTAAAGAGTGAAMAAACAC 2372
2299 CGATGAGTGTGTGCTAGAGACTGTGATTCAGTGTCTTAAATGGGTTAAAGACCC 2358
2373 GGGAGAGGAGCAAGTGAAGGAAAAAGTATCTCATCTGCTAAACGGGTGTGTCGTC 2432
2359 GTTAACACTCTGTATCATTTGATGAGCATTTGCTCCATGACGAGGACGCTGCGCA 2418
2433 GTGACATCTTATATGTGAGAGAGGCTTTGCTTGCATTCGGGATCTCTGCTGCTTA 2492
2419 ATTGCTATCTGCAAACT---AAGAAAGTGTATTTGTGCGGAGACCAAAACATGCG 2475
2493 ATTGCTCTTTTAAACCTCGAGCAAAAGTGTGTATGTGCGAGACCCCAACGAATGCG 2552

2476 TTCTTAACATGATGTGCTGAAGATGATATTTTAACCATGATATGACGTGATAC 2535
2553 TTCTTAATATGATGATGAGCTTTAAGTGAATCTTCAACCAACAACTGACATGAATGT 2612
2536 CATTAAGCATCTCTAGAGGTGACACAGACTGTAAACCGCATGCTCTCAAGCTTTC 2595
2613 CATTAAGTATATCAAGACCTTGCACCGCTCAAGTCAAGCATGCTGTACCTTCAC 2672
2596 TACGACAAGGAATGAAGAGGTTAAACCAATGTGTGATTAATCATATGATACACA 2655
2673 TACGAGGCAAGATGCGACGACCAACCCGTGCAACAAACCATTAATCATAGACACA 2732
2656 GGGACCAAAAGCGGCAAAAGATGATCTGATTTAACTGTTTCAGAGAGTGGGTAAA 2715
2733 GAGACAGCAAGCCCAAGCCAGAGATGATGTTTAACTGTTTCGAGGCTGGGCAAG 2792
2716 CAGTACAGATTGACTTCAAAATACGAAATCAATGCTGCGGTGATGCAAGACTT 2775
2793 CAGCTGAGTTGACTACCTGACACAGAGTATGACAGACGATCTCAGGGCTTC 2852
2776 ACCGGAAAGCGCTTTATGCTGTCAAGTCAAGAGTCAAGGATCACTCTGCGAG 2835
2853 ACCGCAAGGGGTATAGCGCTGTAAGAGTGAATGAATATCTGTATGCCCC 2912
2836 ACTTCTGACGACGTGAACGTGTTACTTACAGCAGCAAGAAAAAGCATTTCTGAGAC 2895
2913 GCGTGGAGCAGCTGAATGTACTGTGACGCGCATGAGATAGGCTGATGTGGAACG 2972
2896 CTAGCTGTGATCCCTGATTAAGACATTAACACTTAATATCCGGGATTTCAAGCT 2955
2973 CTGCGCGGCATCTCTGATTAAGTCTTCAAACTTCCACAGGGTAACTTTACGCG 3032
2956 TCAATGAGACATGCGACGCGGCAACAGAGCCATTATGGCAGCGCTTGTGTAACCG 3015
3033 ACATTGAGAAATGCGCAAGAAACAGACAAATTAATGAAGTATGAAGACCGGCT 3092
3016 CAGACAGCTATGTTTCCAGAAATGAAGTGAAGCTGTGCGGCGAAGGCTTTAGACCA 3075
3093 GCGCTGTGAGCGGTTCCAGAAACAAAGCAAGTGTGTGGGAAAGCCGTGTGCT 3152
3076 GTCTTGGCCACGCGCAACATTGTGCTGACGAGACAGCATGGGAGAC--GTTGACCCA 3132
3153 GTCTGAGACCTGCGGAATCAGATTGACAGAGAGTGTGAGACCATTAATTAACGCA 3212
3133 TTCAAGCATACAGAGCTACTGACCTGAATGGAATGGAATCTTTTGCACACAGGTT 3192
3213 TTTAAGAGACAGAGTACTCTCCAGTGTGCTGTGATGAATTAATTTTCCACCAAGTAC 3272
3193 TTGAGATGACCTGACAGTGTGATTTTTCGCTCTTACCGTGCCTTACTTACAGG 3252
3273 TATGAGTGTACCTGACAGATGGCGTGTTTTGTGCCCCGAAGGTGCTGTATTAAGAG 3332
3253 GATCAGACTGGGATTAATCTGCCAGGAGAAACATGTATGGGCTTAAATGAAGAGTGA 3312
3333 AACCAACACTGGGATTAACAGACTGTGGAAGATGTATGATTCATATGCGCAACGCT 3392
3313 AAGAGTGTCAAGCGGATATCGGTGATCAAAAAGCGGTTGACACAGGAGGTAGCT 3372
3393 GCCAGCTGAGAGTGAACATCTTCTGAAGGGGACGTGGAATGAGGCAAGGCA 3452
3373 GATTAAGAAATTAATCATCAAGACTACTCTCAACAAATTAATGTGTTTCAATTAAT 3432
3453 GTTATGCGAAGAAAGAAATTCACACGCTTTGTGTGTGACATATTAATTCATATAC 3512
3433 CGCGGTTGCCCACTCTGTTGATGTTGACCAACAAAGACAGGATCAATGATCACAGC 3492
3513 CGCAGCTGCGGACCGCTGTGTGCTGATGATCAAGACGTTAAAGCAATGAGGTTGAG 3572
3493 GAATCTCTATTAAGATGAAGAGCAATCTGTGTGTGATTCGCGCATCTATCAGATT 3552
3573 TGTGTGTATTAAGTAAGTAAGGATACCACTGCTGTGTGTGTGTGTGTGTGTGTGT 3632
3553 CAGAGGAAGAAATGAGTCCATGGTCACTTGCCTCACTAATAC-----CATCAGG 3603

Db 3633 TTGCTCGACGAGGATCACTTGTGTACCGCTGAATGTGACACGCGCCGATAGTGC 3692
Qy 3604 TGTGATCTCGATTTGGGAATACCTAGCCATGTCCGTAATATGACATATCTTTGCAAT 3663
Db 3693 TACGACCTTAAGTTTAGACTGTCCGGCTGACGCCGGACAGTTGACTTGTCTTTGTGAAC 3752
Qy 3664 GTTAGAACCCCTAGACGAACCATCATCTAACAGTGGAGATATCCGCTATCCACAC 3723
Db 3753 ATTACACGGGAATTCAGAAATCCACCACTACAGAGTGTGTGACACAGCCATGAAAGCTG 3812
Qy 3724 AGCATGCTTAAGCTGTAAAGCTGTCCACCACTGAAACCTGGGAAACATGTGTGCTATA 3783
Db 3813 CAGATGCTTGGGGAGATGCGCTACGACTGTAAACCCGGCGGCACTCT--TGAATGAGA 3869
Qy 3784 GGGTATGGGCTTGTGATGGGCAACCGGAATATCATCTGCGGTGGCAAGCTCATATT 3843
Db 3870 GCTTACGGATACCGCATTAATAATACAGCCGAGCCGTTGTCTCTTAAGCAAAAGTTC 3929
Qy 3844 AGGTTTACCCCTGTGTCTGCTAGCCCTAAGAACACTGGCCGAATATGAGGTTCTTTGCTG 3903
Db 3930 TCGTGTGACAGAGTGTGTGCGCCGATGTGTGTACACGAAATACAGAAAGTCTTTGCTG 3989
Qy 3904 TTCTTGGGACGAACAACGGCAACACACATGACACAGACAGACTGCGTGTAGTCTT 3963
Db 3990 TTCTCAACTTTGACAAACGGAAGAGACCTCTACGCTACACAGATGAAATACCAAGCTG 4049
Qy 3964 GACAACTCTATCAAGGGCTCAACCGATACAGAGGAGGAGAGCTCCAGCCATACAGATG 4023
Db 4050 AGTGCCGTGTATCCGGAGAACCAATGACACAGCGCGGGTGTGACCACTCTACAGAGAT 4109
Qy 4024 ATCAGAGGTGACATTAGCAAGAGCCCTGACCAAGCTATGCTTAATGCTGCTAATAGCAAA 4083
Db 4110 AAGAAAGGACGATATAGCCACGTCGACAGAAAGGCGTGTGTAAAGCAGCTAACCCCGT 4169
Qy 4084 GGTCAACCAAGTTCCGAGTGTGCGGTGACCTGTACCGAAATGCGCGGCTGCTTTGAT 4143
Db 4170 GGAACCTGTAGGGGATGCGGTATGACAGGCGCTGGCGAAGAAATGCGCCGACCTTTAG 4229
Qy 4144 AGACAGCCCAATGCTGTGGGAGCGGCTAGACTTTGGAAGCAAGAACCG---CTCATCATTA 4200
Db 4230 GGAGCAGCAACACAGTGGGCAAAATTAACAGTCAATGTGTGGCTGTGACCCGCTCATTC 4289
Qy 4201 CATGCTGTAGACCCCAATTTTCTAAGATGCCGGAACCGAGGGGACCTTAAGCTGCGA 4260
Db 4290 CACGCTGTATGCGCTTAATTTCTCTGCGACGATGAAAGCGGAAGCGCAATTTGGCC 4349
Qy 4261 GCTGCTTACATGAGCATAGCGGTTCATTCGTCAACGCTGAGCGGATTAACAAAATATCAGTA 4320
Db 4350 GCTGCTACCGGGCAGTGGCCGCCGAAGTAAACAGACTGTCTCATGAGCAGCGTACCATTC 4409
Qy 4321 CCGCTACTGTCAACCGGCACTTATTTGTGTGGCAAAAGATGAGTGAATCATTTGAT 4380
Db 4410 CCGCTGCTGTCCACAGAGTGTTCAGCGCGGAGAGTAAAGGCTCAGAAATTCCTCTAAC 4469
Qy 4381 CACCTGTTCACTGCTTTGACACTACGAGTGGCGATGTCAGATCATATTTGCTTGATTA 4440
Db 4470 CATCTATTACAGCAATGAGCCCAAGGAGCTGACGTGACCATCTACTGCAAGAACAA 4529
Qy 4441 CAATGGAGAACCAAGATATCGAGGCACTTCCGCCAAGAAAGGCTCGAAATTTGAT 4500
Db 4530 AGTTGGAGAGAAATTCAGAGAACATGTACATGAGAGCGGCTGTGAGTGTCTCAT 4589
Qy 4501 GATGCAAGCCAGTATGACATTTGCTGAGGGTTCACCCAAACAGCTCTTTGGCAGGC 4560
Db 4590 GATGACGTGAGACTACCAACAGACTTGTGAGAGTGCACCCGAGACGACCTGTGTGGT 4649
Qy 4561 AGACACAGTACTCCGTCAATGAGGCAAGTTGTATTAATCTAGGAAAGTACAGATTC 4620
Db 4650 CGTAAGGGCTACAGTACCACTACCGGGTGTGTACTGTCTTGAAGATGCAAAATTC 4709
Qy 4621 CATGACCGCCAGAGCAATTCGCAAAATTCATGCAATGTGGCCCAACAAATCTGAGCT 4680

Db 4710 AACGAGCTGCTATTGATATGACAGATATGACGTTGTGGCCAGATCGACAGAGCA 4769
Qy 4681 AATGACGATTTGCTGTATGACATCTGGGGGAGATATGTCCAGCATCGGCTCCAATGC 4740
Db 4770 AACGACAGATATGCTTATGAGCGCTGGCGGAACATGACACATCATGATCCAAATGT 4829
Qy 4741 CCAGTAGAGAGTCAAGAGCGCTGTGCTTCACCTCAACACACTTCCATGCTGTGTAATAC 4800
Db 4830 CCGGTGAACGATTCGATTCATCAACACCTCCAGAGCAGTGCCTGCTGTGGCGTAC 4889
Qy 4801 GCTATGAGCGCTGACCGGCTATACAGTTGCGCTTGGCAAGAAAGAACAGTCCCGTA 4860
Db 4890 GCAATGACAGCAGAACGATGCGCCCTTGTGTCACACCAAGTTAAAGCATGTGTGTT 4949
Qy 4861 TGTCTATATTCCTGTGGCGGAGTACAGATTCACAGCGGTGACAGACTACGTGACGC 4920
Db 4950 TGTCTATTTTTCCTCTCCGAATATCATATGATGAGGGGTGCACAAAGTAAAGTGCAG 5009
Qy 4921 AACCAAGTCCGTTTTCAGGCGGTGTACACCGGCTGTACACCCGAGAAATGACCGGAA 4980
Db 5010 AAGTTCTCTGTTGACCGACCGTACCTTCAAGTGTTAATGTCGGGAAATATGCCCA 5069

RESULT 9
AAC5473
ID AAC5473 standard; DNA; 12464 BP.
XX
AC AAC5473;
XX
DT 11-JAN-2001 (first entry)
XX
DE Destination vector pDEST9 nucleotide sequence.
XX
KM Bacteriophage lambda; att; recombination site; attB; attP; attR; attL;
KM mutant; recombinational cloning; entry vector; destination vector;
KM gene product targeting; fusion tag cleavage; de.
XX
OS Bacteriophage lambda.
OS Synthetic.
XX
PN WO200052027-A1.
XX
PD 08-SEP-2000.
XX
PI 02-MAR-2000; 2000WO-US05432.
XX
PF 02-MAR-1999; 99US-0122389.
XX
PR 23-MAR-1999; 99US-0126049.
XX
PR 28-MAY-1999; 99US-0136744.
XX
PA (LIFE-) LIFE TECHNOLOGIES INC.
XX
PI Hatley JL, Braach MA, Temple GF, Cheo D;
XX
DR WPI; 2000-543948/49.
XX
PT Isolated nucleic acid molecules encoding an attB1, attB2, attP1, attP2,
XX
PT attL1, attL2, attR1, and attR2 nucleotide sequence useful for the
XX
PT recombinational cloning of polypeptides -
XX
PS Disclosure; Fig 29; 459pp; English.
XX
CC The present invention describes isolated nucleic acid molecules (I)
CC encoding an attB1, attB2, attP1, attP2, attL1, attL2, attR1, and attR2
CC nucleotide sequence. Also described are: (1) an isolated nucleic acid
CC molecule (II) comprising one or more att recombination sites comprising
CC at least one mutation in its core region that increases the specificity
CC of interaction between the recombination site and a second att
CC recombination site; and (2) an isolated nucleic acid molecule (III)
CC comprising one or more mutated att recombination sites comprising at
CC least one mutation in its core region that enhances the efficiency of
CC recombination between a first nucleic acid molecule comprising the
CC mutated att recombination site and a second nucleic acid molecule

comprising a second recombination site that interacts with the mutated
act recombination site. (I), (II), (III), primers, vectors and methods
from the present invention are used for the recombinational cloning of
nucleic acid molecules. They can be used for changing vectors, targeting
gene products to intracellular locations, cleaving fusion tags from
desired proteins, operably linking nucleic acid molecules of interest to
regulatory genetic sequences, constructing genes for fusion proteins,
changing copy number, changing replicons, cloning into phages and
cloning. (I), (II), (III), host cells and vectors can be used in the
production of polypeptides and antibodies. The present sequence is
used in the exemplification of the present invention.

Sequence 12464 BP; 3328 A; 3120 C; 3257 G; 2759 T; 0 other;

Query Match 31.3%; Score 1565.6; DB 21; Length 12464;
Best Local Similarity 58.3%; Pred. No. 0;
Matches 2905; Conservative 0; Mismatches 2039; Indels 36; Gaps 8;
QY 31 AGAATTCACGTTGATGATGCTGACAGCCGATGTCAGTGGTTACAGCGGACGTT 90
DB 5361 AAAGTCATGTTGATTTAGAGCTGACAGCCATTCATCAAGTCTTTGACAGAGCATTT 5420
QY 91 CCACAAATTTGATGATGAGCAAGGCAAGTCACTGACAAATGACATGCGCAATGCCAGAGCG 150
DB 5421 CCGTCGTTGAGGTGAGATCATTCAGAGTCAACCAATGACATGCAATATGCCAGAGCA 5480
QY 151 TTTTTCGATGTCGCAACAAAGCTCATTTGAGCGCAAGTCCAGCGGACCAAGTTATCTTG 210
DB 5481 TTTTTCGACCTGGCTTACCAATTTATCGAGGAGACATGCAAAAGACACATCATCTTG 5540
QY 211 GACATTTGGAAGTGGCCCGTCAACATGACATTCATTCACCGCTATCATTTGATCTGC 270
DB 5541 GATATCGCAGTGGCCCTTCACAGAGATGATGTCAGCAAAATACCATCTGCGTATGC 5600
QY 271 CTATGATAGCGCTGAAGACCCGACAGACTCAACGCGTATGCAAGAACT-----T 324
DB 5601 CCTATGGCAGCGGAGAAAGACCCGAAAGGCTCATGATCTAGCAAAAGAACTGGCAGCG 5660
QY 325 AAGAAAGTGATCATTTACGAAAGAACTTACCTCTAAGGGGAGACCTCTGMAAGTC 384
DB 5661 GCCTCCGGAAGGTCTGAGATGAGAGATCGAGAAAATACCGGACCTGCAACGCTC 5720
QY 385 ATGTCAACACGAGACGAGACTCATCTCTGTATGACACAGAGCCAGCTGTAG 444
DB 5721 ATGGCTACGCAAGACGCTGATCTCTACCTTTGCTGCACTACAGAGTACGCTGCT 5780
QY 445 TACTTTGAGAGTGAAGCATATACCAAGATGTATCGAGTCCATGACCGACATCAATC 504
DB 5781 ACGGACGCCGAAGTGGCGGTATACGAGACGCTGATGCTGATGACCAATCGCTG 5840
QY 505 TACCAACAGGCGCTTAAAGAGTTAGACAAATTTACTGATAGGCTTTGACAGACCTT 564
DB 5841 TACCATACAGGAGTGAAGAGGTGTCAGAACGCGATGAGTTGGTTTGACACACCCG 5900
QY 565 TTTATGTACAAATAATGAGGCTTCTACCTTACACAGAGTGGGCTGACAG 624
DB 5901 TTTATGTGACGCGCTAGCAGGCGGCTATCAACCTACGCAAACTGGGCGAGAG 5960
QY 625 AGAGTATTTGAGAGCAGTACATTTGCTCGTAACTTCAGATCTTCAGAGAGAGCGTT 684
DB 5961 CAGGTGTTACAGGCGAGGACATAGAGCTGTGAGAGCATCTTGATGAGGAGAAATC 6020
QY 685 GGAAGAACTCTCAATCTTGAAGAGAGGCTCCAACTACTAATAGATCATATTTCTG 744
DB 6021 GGGAAACGTCTCATCTTCGCAAGAAAGCAATTTGAAACCTTGCAACAGTATGTCTCG 6080
QY 745 GTTGGTCAACATCTACAGAGAGATGATCACTGTTACAGTGGCATCTTCAAC 804
DB 6081 GTAGATCTACATTTGATCACTGAGAGCAAGAAAGTACTGAGAGAGTGCATTAACCTCC 6140
QY 805 GTGTTCACCTTGAAGAGAAAGTCTAACTTCAAGGTAGATGTGGACCATTTGAGCTGT 864
DB 6141 GTATTCCACCTGAAGAGTAAACATCTTTACCTGTAGGTGAGATACATCATCATGT 6200

QY 865 GAAGGTAGCTCATCAAAAAGATTAACGATCAGCCAGACTATACGGTAAAGTTGAGAAC 924
DB 6201 GAAGGTAGCTAGTTAAGAAAATACATATGTGCCCCGCTGTACGTTAAACGGTGGG 6260
QY 925 TTGGCGTCAACAAATGACATCCGAGGGTTTCTTGAATTCAAAGTCAACATAGCGTGGC 984
DB 6261 TAGCGGTGACATCAACGAGGAGATTCCTAGTGTGCAAGAACACAGACACTGTCAAA 6320
QY 985 GGGGAGGGTTTCTTTGCTGTGTGTATGATGATACCAACCACTTTGCCATCAGATG 1044
DB 6321 GGAAGAAAGTCTCATCTCCGTATGCACTTACCTGCTCCCAACCATCTGTATCAATG 6380
QY 1045 ACAAGATTCGCACTGACGTTAGTGTGATGACGCAAAAACATTTGTTGGGCTC 1104
DB 6381 ACTGCACTTACGACCGACGCTACACCGGAGAGCAGAAAGTTGTTAGTGGATTTG 6440
QY 1105 AACCAAGATTTGCTGCAATGATGAGACGCAAAATACTAACACATGACAACTAT 1164
DB 6441 AATCAGAGATAGTTGTGAAAGGAGAAACACAGGAAACATACAGATGAAGAACTAT 6500
QY 1165 CTATTAACAGTGTGCTCCAGGGCTTTCCAGGTGGCGGTGAACATGCTGCCACTTG 1224
DB 6501 CTGCTTCGATTTGTGGCGCTGCAATTTAGCAAGTGGCGGAGAAATCAAGCAGACTT 6560
QY 1225 GAGGAGAGAAAGAACTAGGGGTGGGAGGCGCACTTCTATGAGGCTGCTGCGGCT 1284
DB 6561 GATGATAAATAAACCCTGAGGTGTCGAGAGAGTCACTTACTTGTCTGTGTGGGCA 6620
QY 1285 TTCAAGACCCGAAATACATCATCTACAAAGAGCTGTACGCAAAATTAAGAAA 1344
DB 6621 TTTTAAACGAGAAAGATGACACATGTACAAAGAACACACGCAAAATAGTAGAG 6680
QY 1345 GTACTTCGCTCTTTGATCTATTTGTATTCACGCTTACAGCCAGGCTGTGATAG 1404
DB 6681 GTGCTTCAAGATTTTAACTGTGTCTATCCGAGCTTATGCTTACAGGCTCTGCAATC 6740
QY 1405 GCGTTCGCGGTGAGCTCAAGCTGCTGTGAACCAACTGTCAACCCGACCGGCTATT 1464
DB 6741 CCACTCAGATCAGCATTAAGATGCTTTTGGCCAAAGAACCAAGC---GAGATTAA 6797
QY 1465 ACAATGCGCATGTGAGCATCTGCGTGTCTTACAGCAAGAGCTGAAGAGTGGCTCA 1524
DB 6798 CCGTCTCGACGCGGTGTGACGAGGAGATGTGAACAAAGAGGAGAGAGTGGAG 6857
QY 1525 GCGGAAGATCAGAGAAAGCCCTGCAACCTTGTCTCCCTGAATAGAAAAGAGACCTGA 1584
DB 6858 GCCGAGCTGACTAGAGAGGCTTACCAACCCCTGCTCCCATCGCGCGGAGAGCGGA 6917
QY 1585 GAGGAGAGTGAAGCTCATTTATGCAAGAG-----GCAGAGAGGTAGGTGAGACA 1638
DB 6918 GTCTGCACTGCACTGTTGAAGAACTAGATATCAGCAGAGTGCAGGGGTGAGAAACA 6977
QY 1639 CCAAGAGACATCAGGCTGAGCAAGTTTACCGAGCGAAGAGAAAGATTTGGTCTTACGCT 1698
DB 6978 CCTCGACGCGGTTGAAGATCACCGACAGCCGAAGCAGTACTACGAGAAATTAAGTGA 7037
QY 1699 ATACTTCAACCCAGGCGGTATGATATGAAAACTGGCGGTATACCACTTGGCG 1758
DB 7038 GTTCTGTCCCGAGAGCCTGTCAAGAGCTCAAGATTTGCCCCCGGTGACACCTCTGCA 7097
QY 1759 GAACAAGTACTGGAATAGATCAACAAAGTATGAGGCAAGGATACAAAGTGCAGGCTAC 1818
DB 7098 GAGCAGGTGAATTAATTAACATTAAGGAGAGGCGCGGTTTACAGGTGACGAGTAT 7157
QY 1819 CACGTAAGGTCAATTTGACAGAGAGGAGCGCGGTCCCTGTTCACAGACTTCAAGCAATG 1878
DB 7158 GAGCGCAGGCTCTACTACATGATGATGCGGCAATTCGGATCCCTGATTTCAAGCTTGG 7217
QY 1879 AGTGAAGCGCTGATGCTTTTCAACGAGAGGAGTTGCTAAACAGATACCTGCACAC 1938
DB 7218 AGCAGAGGCGCACTATGTGTATCAACGAAAGGAGTTGTCTCAACGAAATTAATACAT 7277

1939 ATGCAATCAACGAGAGCGCTAAACATGACGAAAGTACTATATAGACTGTAAAGACT 1998
1939 ATGCAATCAACGAGAGCGCTAAACATGACGAAAGTACTATATAGACTGTAAAGACT 1998
7278 ATTGCGTTTCAACGAGCCCTGCTGAACACCGACGAGGAACTTAGAAGAACTCAGAGCT 7337
1999 CAGGACACAGACTCAGAAATAGCTTTCCATATTGACGACGAAAGTGTATTAGCGAGA 2058
7338 GAAAGAACTGACGCGAGTACGCTGTTCGACGTAGATTAATAAAATGCTGCTCAAGAGAG 7397
2059 GACGAGGTCCTTGTGCTTAACCGGTATCTGTAGATTCACATTCACAGAGTTGG 2118
7398 GAAAGCTGGGTTTGTGTGTGTGAGAGCTAACCAACCCCGTTCCATTAATTCGCG 7457
2119 TACGAGATCTCAAGACAGACGACGACCTCAAAAGTCCCAACCATCGAGTCTAT 2178
7458 TACGAGAGGCTGAATATAGAGCGCTCGGACCATTAAGACTACAGTAGTAGAGTCTTT 7517
2179 GAGATGCCAGTTCAAGTAAATCTGGAATCATCAAAAGCGCTGTGACTTAAAGAGATCTG 2238
7518 GGGGTTCCGGATCAGGCAAGTCTCTATTATTAAGAGCTGTGACCAACACGATCTG 7577
2239 GTTGTAGTGCAGAAAGGAAATCGCGCAAAATCATGAGGATGTAAAGAGATGAGA 2298
7578 GTACACGCGGCAAGAGAGAGACTGCGAGAAATATGTTACAGCTGAAGAGACCGC 7637
2299 CGTATGATGTTGCTGTAGGACTGTGATTCAGTCTTCTAAATGGGGTTAAGACCCC 2358
7638 GGGAAAGGGAACAAGTAGGAAACAGTACTCTCTCTAAAGGGGTGTGTGTGCTC 7697
2359 GTTAACACTCTGTACATTGATGAGGCAATTTGCTGCTCATGAGGACGCTGTGACCTG 2418
7698 GTGACATCTATATGTGACGAGGCTTCTGCTCAATTCGGTATCTGCTGCGCTTA 7757
2419 ATTGCACTGCTCAAACT--AAGAAAGGTATGTGGGGGAGCCCAAAACAAATGCGC 2475
7758 ATTGCTCTTTTAAACCTGAGCAAGGTGTATGTGAGAGCCCAAGCAATGCGGA 7817
2476 TTCTTTAATCATGATGCTGCTGAAGTAACTTTTAAACATGACATATGACTGAAGTAC 2535
7818 TTCTTCAATATGATGACTGATTAAGTGAATTTCAACCAACATCTGCACTGAATATAT 7877
2536 CATAAAGCACTCTTAGAGAGTGCAACAGACTGTAAACCGCATGCTGTCTCAACGCTCTTC 2595
7878 CATAAAGTATATCCAGAGCTTGCAACGCGTCCAGTCCAGCATGCTGTCTCACTTGCAC 7937
2596 TACGACAGCGAATGAAGCGTTAACCCATGCTGTATTAATATCATATGATATCCACA 2655
7938 TACGAGGCAAGATGCGACGACCAACCCGTGCAACCAACCATATATCATAGACACACA 7997
2656 GGGACCAAAAGCCGCAAAAGATGATCTGATTTCTAACCTGTTTCAAGAGATGGGTGAAA 2715
7998 GGACAGACCAAGCCCAAGCAGAGACATCGTGTAACTGCTTCCGAGGCTGGGCAAG 8057
2716 CAGCTACAGATTGACTACAAAATATCAGAATATCATGACTGCGGTGCAATCGCAAGACTT 2775
8058 CAGCTGAGTTGACTACCGTGACACGAAGTCAAGACGAGCACTTCAGGCGCTC 8117
2776 ACGGGAAGGCGTTATGCTGTCAAGTAAAGTCAAGAGAAATCCACTCTACTCGAG 2835
8118 ACCGCAAAAGGGATATGCGCGTAAAGGCAAGAGTGAATGAATATCCCTGTATGCCCCCT 8177
2836 ACTTCTGACAGCTGAACGCTGTACTTACACGACAGAAAAAGCATTTGTCTGAAGAG 2895
8178 GCGTGGAGCACTGAATGTACTGTGACGCGCACTGAGATAGGCTGTGTGGAAAAAG 8237
2896 CTAAGTGTGATCCCTGTGATTAAGACCTTACAGCTAAATATCCCGGGGATTTCCAGGCT 2955
8238 CTGGCGCGGATCCCTGATTTAAGGTCTATCAAACTTCACAGGATTAATTTTACGCGC 8297
2956 TCATTGAGAGCTGCGACGCGCAACAGACGCAATTAAGGACGCGCTTCTGATTAAGCGC 3015
8298 ACATTGAGAGATGGCAAGAAAGACAGCAAAATATATGAGTATTAAGAGACCGGCT 8357
3016 CAGACAGCTGATGTCTCCAGATTAAGGTGAAGTCTGCTGGGCAAGGCTTTAGAGCA 3075

8358 GCGCTGTGAGACGCGTTCAGAAACAAAGCAAGCTGTGTGGGCAAAAAGCTGTGCTC 8417
3076 GTCTTGGCAGCGCCCAATTTGTGTGACGAGACAGCTGGAGAC--GTTCACCCA 3132
8418 GTCTTGGCAGCTGCGGATCAGATTAACACAGAGGAGTGAAGCACAATATTAACAGA 8477
3133 TTCAAGCATGACAGAGCTACTACCTGAATGAGCACTGAACTTTTGTGACGAGTTTC 3192
8478 TTTAAGAGGACAGAGCTTACTCTCAAGTGTGCTGAAATGAATTTTGCACAAATAC 8537
3193 TTTGAGTAGACCTGAGCAGTGGTATTTTCCGCTCTTACCGTGCACCTTACTTACAG 3252
8538 TATGAGTGAACCTGAGCAGTGGCTGTTTCTGCGCCGAAGTGTCCCTGATTAACAG 8597
3253 GATCAGCACTGGGATTAATCCGCAAGGAAACATGTATGGCTTAATAGAGAGTAGCA 3312
8598 AACAACTGCGGATTAACAGACTGTGTGAAGAGATGTGAATTCAAATGCGCAACAGT 8657
3313 AAGAGTGTGACGCGGATATCCGTCATCAAAAGGCGTTGACAGAGCGAGGTAGCT 3372
8658 GCCAGCTGGAAGCTAGACATATCTTCTGAAGGGGACGTGCAATCGGCAAGAGCA 8717
3373 GATATTAAGAAATTAATCCATCAAGGACTACTCTCCAAATTAATGTGTTCCATTAAT 3432
8718 GTTATCGCAAGAAAGAAATCCAAACCGCTTCTGTGCTGAGCAATGATTAATCTTCAAC 8777
3433 CGCGGTTGCCCCACTGCTTATGCTTGAACCAAGAGCAAGGATCAACTGATACAGC 3492
8778 CGCAGGCTGCGCACGCGCTGTGTGAGTGAAGAGCGTTAAAGGACAGTAGGTTGAG 8837
3493 GGAATTCATATTAAGTGAAGGCAAAATCTGTGTGTGATCGGAGCTCATACGANT 3552
8838 TGGCTGCTAATTAAGTGAAGGATCCAGCTCTGCTGTGATGATACCACTGCTC 8897
3553 CAGGGAAGAAATGAGATGATGCTGATGCTATTCGCTCAATTAAC-----CATCAG 3603
8898 TTGCTGAGCAAGGATCACTTGTGTGTCACCGGTGAATGTCACAGGCGCGGATAGTGC 8957
3604 TGTGATCTGATTTGGGAATACCTAGCCATGTCCGTAATATATGATATTTCTTTGCAAT 3663
8958 TACGACTTAATTAAGACTCCGCTGACCCCGCAGAGTTGCACTTGTGTTGTGAAC 9017
3664 GTTAGAGACCCGCTACAGAACATCACTACCAAGAGCGAGATCAGCTATCCACAC 3723
9018 ATTACACGGAATTCAGATCACCATCACAGACGATGTGACACGACCATGAAAGCTG 9077
3724 AGCATGCTAAAGTGAAGGCTGTCCACCACTGAAACATGTCGCTATTA 3783
9078 CAGATGCTTGGGGAGAGTGGCTTAAGACTGCTTAAACCCGCGGATCT--TATAGAGA 9134
3784 GGGTATGGGCTTGTGATCGCGCAACGGAATATCATCTGCGGTGACAGCTCATTT 3843
9135 GCTTACGATACGCGGATTAATACAGCGAACCGTGTCTCTTAAAGCAAAAGTTTC 9194
3844 AGTTTACCGGTGTCTGACCTTAAGAACACTCCGAAATATCTGAGTCTCTCTGCTG 3903
9195 TGTGTGAAGAGTGTGCGCCGATTTGTCTACACGAAATACAGAAATGTTCTTGTGCT 9254
3904 TTCTTGGCAAGCAACGCGCAACACACATGACAGAGCACTGCTGTAGTGTCTT 3963
9255 TTCTTCAACTTTGACACAGGAAAGAGACCTCTTACGATACACAGATTAATCAAGCTG 9314
9315 AGTCCGATGTATGCGGAGAGACCATGACACGCGCGGATGTGACCATCTTACAGAGTT 9374
4024 ATCAGAGTGAATTTAGCAAGAGCTGACCAAGCTATCTGTTATGCTGTAATAGCAA 4083
9375 AAGAGAGCAATAGCCACGTCACAGAACGCTGTGTTAAACGACGCTAACGCCCT 9434
4084 GGTAAACAGGTTCCGAGTGTGCGGTGACCTGTACGAAATGCGCGGCTGCTTTGAT 4143

```

Dh 9435 GGAACGTGAGGAGATGGGTATGACAGGCGCTGGCGAAGAAATGGCCGTACCTTTAAG 9494
Qy 4144 AGACAGCAATAGCTGTGCGGACGGCTAGACTTGTGAAGCAGAACG--CTATCATTA 4200
Dh 9495 GGACAGCAACACACAGTGGGACAAATTAAACAGTCAATGTGGCTGTACCCCGTCAATC 9554
Qy 4201 CATCTGTAGAGACCAATTTTCTAAGATGCGGGAACCGGAGGCGACCTTAAGCTGCA 4260
Dh 9555 CAGCTGTAGGCGCTAATTTTCTGTCCACGACTGAAGGAGGAGGACCGGAATTGGCC 9614
Qy 4261 GCTGCTACATGAGCATAGCTGCATGCTCAAGCCTGAGCGATTACAAAATATCAATA 4320
Dh 9615 GCTGTACCGGCGGACGTGGCCGCGAAGTAACAGACTGTCACTGAGAGAGGTAGCATC 9674
Qy 4321 CCGTACTGTCAACCGGATCTATTGTGTGGCAAAATGCAAGTGAATCATTTGAT 4380
Dh 9675 CCGTGTGTGTCAAGAGATGTTCAAGCGCGGAAAGATAGGCTGACAGCAATCCCTCAAC 9734
Qy 4381 CACCTGTCACTGCTTTGCACTACGATCCGATGTCACCATATATTTGCTGATATA 4440
Dh 9735 CATCTATTCAAGCAATGAGACGCAACGACGCTGACGTCATCTACTGACAGACAAA 9794
Qy 4441 CAAATGGAGACCAAGATTAATGAGGCAATTCACCGCAAAAGAACGCTGCAAAATTCGAT 4500
Dh 9795 AGTTGGGAGAAAGAAATTCAGGAAGCCATTGACATGAGGACGGCTGTGAGTTGCTCAAT 9854
Qy 4501 GATACCAAGCCAGATAGATGATGCTTGTGCTCAGGGTCCACCCCAACAGTCTTTGGCAGC 4560
Dh 9855 GATACCTGTGAGCTGACCAACGACTTGTGTGAGTGCACCGGACAGAGAGCTGTGGGT 9914
Qy 4561 AGACCAAGTACTCCGTCAATAGGAGGCAAGTTGATTCATCTCGAAGATACAGATTC 4620
Dh 9915 CGTAAAGGCTACATGACCTGACGCGGTGCTGTACTGTGAAAGATACCAAAATTC 9974
Qy 4621 CATCAGACCGCAAGACATTTGCGAAATCCATGCAATGTGGCCCAACAAATCTGAGGCT 4680
Dh 9975 AACCAAGCTGTATTTGATATGAGAGATGACAGATGCTGATGTGGCCAGACTGCAAGAGCA 10034
Qy 4681 AATAGCAAGATTTGCTGTATATCTGTGGGAGAGATATGTCCAGATTCCTCCAAATGC 4740
Dh 10035 AACCAAGATATCTATATGAGCTGTGGCGCAAAACAAATGCAACATCAAGTCCAAATGT 10094
Qy 4741 CCATGAGGAGATGAGAGGCGTGTCTGCTCACTCAACACTTCATGCTGTGTAATATC 4800
Dh 10095 CCGGTGAACATTCGATTCATCAACACTCCAGAGACAGTCCCTGCTGTGCGCTAC 10154
Qy 4801 GCTATGACGCGTGAACGCGTATACAGTTGCGCTCTGCGAAGAAAGAACAGTTGCGCGTA 4860
Dh 10155 GCAATGACAGCAGAACGATGCGCGCTTGTAGTCAACCAAGTTAAAGATGATGTT 10214
Qy 4861 TGCATCATCTCTGCTGTGCGGAATGACAGATCAACAGCGCTGCGAAGATCACTGACG 4920
Dh 10215 TGCATCATCTTTTCCCTCCCAAAATACATGTAAGTGGGTCGAGAAAGTAAAGTGGCAG 10274
Qy 4921 AAACAGAGCTGTTTTCAGGCGGTGTACACCGGCTGTACACCCAGAGAAATACGCGAA 4980
Dh 10275 AAGGTTCTCTGTTGACACCGCAGCGTACCTTCAAGTGTAGTCCGCGAAGATATGCCGA 10334

```

```

PN CN1367020-A.
XX
PD 04-SEP-2002.
XX
PF 19-OCT-2001; 2001CN-0128979.
XX
PR 19-OCT-2001; 2001CN-0128979.
XX
PA (VUHU-) UNIV HUIXI HOSPITAL SICHUAN.
XX
PI Zhao L, Qin S;
XX
DR WPI; 2003-168348/17.
XX
PT Nucleic acid vaccine capable of being high-effectively and safely
PT expressed comprises a conventional nucleic acid vaccine plasmid and
PT alpha virus replication unit NS1-NS4 -
XX
PS Example; Fig 1; 14pp; Chinese.
XX
CC The present invention describes a nucleic acid vaccine which can be
CC effectively safely expressed. The nucleic acid vaccine plasmid forming
CC the nucleic acid vaccine is formed from a conventional nucleic acid
CC vaccine plasmid and an alphavirus replication unit NS1-NS4 gene
CC recombination. After having been injected into body, under the action
CC of a CMV promoter, the nucleic acid vaccine can transcribe the NS1-NS4
CC replication unit and RNA of its downstream target gene with high
CC replication, and it possesses high expression efficiency of its target
CC antigen, and after its transcription in the cell and expression level
CC has reach a certain high level, it can induce the cell transfected by
CC performing a death process to form a self-destroying mechanism so as to
CC effectively solve the safety problem of nucleic acid vaccines. The
CC present sequence represents an alphavirus replication unit NS1-NS4 gene
CC related nucleotide sequence, which is used in an example from the
CC present invention.
XX
SQ Sequence 7399 BP; 1966 A; 1915 C; 2025 G; 1493 T; 0 other;

```

```

Query Match 31.3%; Score 1564; DB 25; Length 7399;
Best Local Similarity 58.3%; Pred. No. 0;
Matches 2904; Conservative 0; Mismatches 2040; Indels 36; Gaps 8;

```

```

Qy 31 AGAATTCACGCTGATGATGATGCTGACAGCCCGTATGCAAGTGTGTTACACGCGACGTTT 90
Dh 95 AAGATGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 154
Qy 91 CCACAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 150
Dh 155 CCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 214
Qy 151 TTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 210
Dh 215 TTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 274
Qy 211 GACATTTGAAAGTGGCCCGTCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 270
Dh 275 GATATGCGAGTGGCTTCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 334
Qy 271 CTTATGATTAAGCGCTGAAGACCGGACAGACTACACAGCTATGATGATGATGATGATGATGAT 324
Dh 335 CTTATGCGCAGCGAGAGAACCCGAAAGGCTGATGATGATGATGATGATGATGATGATGATGAT 394
Qy 325 AAGAAAGTGAATTAACCGACAGAAACATAGCCCTTAAGCGGAGACCTGTGGAAGTC 384
Dh 395 GCCTCCGGGAAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 454
Qy 385 ATGTCAACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 444
Dh 455 ATGCTACGCGCAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 514
Qy 445 TACTTTGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 504
Dh 515 ACGGACACCGAAGTGGCGGTATACCAAGACGTGATGATGATGATGATGATGATGATGATGATG 574

```

505 TACCACGAGCGCTTAAAGAGTTAGACAATTCTGGATAGGCTTTGACACGACCCCT 564
575 TACCATCAGGGGATGAAGAGTGTCAAGACGGGTAATTGGATTGGTTGACACACCCCG 634
655 TTTATGTACAAAACATGGCAGGCTTCACTCTACTTACACAGCACTGGGCTGACGAG 624
635 TTTATGTTTGAACGGCTAGCAGGCGCGTATCCAACTACGCGCAAACTGGGCGACGAG 694
625 AGAGATTGGAAGCACTGAACATGGCCCTGGTACTGAGACTTCAGAGAGGAGGCTT 684
695 CAGGTTTACAGGCGCAGAGAACATAGGACTGTGTGAGACATCTTCACTGAGGAGGAGCTC 754
685 GAAAACTCTCAATCTTAGAGAAAGAGGCTCAACCTACTAATTAAGTCAATATTCTCG 744
755 GCGAAACGTCTCATCTCCGCAAGAGCAATTGAAACCTTGGACACAGCTATGTCTCG 814
745 GTTGTTCAACATCTACACAGAGATGATCACTGTACGTAAGTGGCATCTTCCAAAC 804
815 GTAGATCTACATGTGTACACTGAGAGAAAGCTACTGAGAGAGCTGGCACTTACCTCC 874
805 GTGTTCCTTGAAGAAAGAGTCTTACAGAGTATGTGGGACCATTTGTACGCTGT 864
875 GTATTCACCTGAAAGGTAAACATCTTTACCTGTAGGTGATACCATGATCATGT 934
865 GAAGGATCGTCAATCAAAAAGATTAAGATCAGCCAGGACTATAGGTTAAGTTAGAAC 924
935 GAAGGATCGTATGTAAAGAAATCACTATGTGCCCCGCTGTACGTTAAACGTTAGAG 994
925 TTGGGTCACAAATGATCGCAGAGGTTCTTGAATTCGAAATGACAGATACGCTGGC 984
995 TACGCGTACGATCAGCGGAGGATTCCTAGTGTGCAAGACACAGACACTGTCAA 1054
985 GCGCAGAGGTTCTTTTGT 1044
1055 GAGAAAGTCTCACTCCCTGTATGACCTACGCTCCCTCAACCATCTGTGATCAATG 1114
1045 ACAGGATTTCTGGCACTGACGTTAGTGTGATGACGACAAAACCTATTGGGCTC 1104
1115 ACTGGCACTATAGCACCAGCTCACCAGAGAGGACAGAACTTGTATGTGGATG 1174
1105 AACCAAGATGT 1164
1175 AATCAGAGATAGT 1234
1165 CTATTCAGAGT 1224
1235 CTGCTTCGATTTGT 1294
1225 GACGACGAGAAAGACTAGGGGTGTGGAGGCACTCTTACTATGGGCTGTGGCT 1284
1295 GATGTGAAACCTCTGGGTGTCCGAGAGGTACTTACTGTGTGTGTGTGTGTGTGTGT 1354
1285 TTCAAGACCCAGAAATCAGATCTCATCTTCAAGAGGCTGTGTGTGTGTGTGTGTGT 1344
1355 TTTAAAGAGAGAGATGACACCATGTATCAAGAAACAGACACCATCAATGTGTGTGT 1414
1345 GTACCTGCGTCTTGT 1404
1415 GTGCTTCAGAGTTTAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1474
1405 GCGTTCCGCGTAGGCTCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1464
1475 CCACTGATATCAGCATTTAAAGTCTTTTGGCAAGAAACCAAGC---GAGATTATTA 1531
1465 ACAATGGCGATGTGAGCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1524
1532 CCGTGTCTGACGGGTGTGACGCGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1591
1555 GCGGAGAGATCAGAGAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1584
1592 GCCGAGCTAGTGAAGAGCCCTTACACCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1651

1585 GAGCAGAGATAGACCTCATTTATGCAAGAG-----GCAGAGCAGTAGCGTGAGACA 1638
1652 GTTCGACACCTGACGTTGAAAGAACTAGAGATACAGAGGTGACAGGGGTGTGTGAACA 1711
1639 CCACGAGACATCAGGGTGAACAAGTTTACCAGCGAAGAGAAAGATTGGTCTTACGCT 1698
1712 CTCGCAAGCGGTGTGAAGTACCGCACAGCCGAACGACGTACTAGAAATTAACGTA 1771
1699 ATACTTTACCCCGAGCGGTATTGAATGTGAAAAAACTGGGTGTATTCACCCATTTGGCG 1758
1772 GTTGTCTCCGACAGACCTGTCTCAAGAGCTCCAAAGTTGGCCCCCGTCCACTCTTACCA 1831
1759 GAACAAGTATGTGTAATGACTCACAAAGTATGGGCGAGGAGATACAAAGTCAAGCCTATC 1818
1832 GAGCAGGTGAATAATATATACATACACGAGAGGCGCGGTTACAGGTGACGAGATAT 1891
1819 CACGTTAAGTCAATTGTATCCAGAAAGGACGGCGGTCCCTGTTCAGAACTTCCAGGCAATG 1878
1892 GACGCAAGGCTCTACTACATGTGTATCGGCAATCCGATCCGATCCCTGAGCTTTCAG 1951
1879 AGTAGAGCGCTAGCATGTGTTCACAGAGAGGAGTGTGTAAACATATACGACAC 1938
1952 AGCAGAGCGCTACTATGTGTGTACACGAAAGGAGATTGTGTCAACAGAACTATACAT 2011
1939 ATCCGAATCAACGGAGGAGCGCTAAACATGACGAAAGATCTATAAGACTGTAAAGACT 1998
2012 ATTGCGCTTACGAGACCGCTGCTGAACCGACGAGAACTACGAAAGTCAAGCT 2071
1999 CAGACACAGACTGAGATACGTCTTGTATTTAGACGACGAAAGTGTATTAGCGAGAA 2058
2072 GAAGAAGTGAACGCGCAGATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2131
2059 GACCGAGTCCCTGT 2118
2132 GAACGCTCGGTTGT 2191
2119 TACGAGATCTCAGACACGACGACGACCTCAACAAAGTCCCAACCATCGAGTCTAT 2178
2192 TACGAGGCTGAAGATCAGGCTCGGACCATATATAGATTAACATAGTATGAGTCTTT 2251
2179 GAGTGCAGGTTCAAGTAAATCTGAAATCATCAAAACGCTGTGATTAAGAAAGTCTG 2238
2252 GGGTCCGGGATCAGGAAAGT 2311
2239 GTTGTGAGTGTGAGAAAGAAATCTGCGCAGAAATCATCAGGATGTATAGAGATGAGA 2298
2312 GTCAACGCGCAAGAGAGAACTGTCCAGGAAATAGTTTACGAGTGTAAAGAACACCGC 2371
2299 CGTATGATGT 2358
2372 GGGAAAGGACAGTATAGGAAACAGTATCTCATCTGTAAACGGGTGTGTGTGTGTGT 2431
2359 GTTAAACTCTGTATCATTTGATGAGGCAATTTGCTGTGTGTGTGTGTGTGTGTGTGTGT 2418
2432 GTGACATCTATATGT 2491
2419 ATTGCCATGTCAAACT---AAGAAAGTGTATTTGTGTGTGTGTGTGTGTGTGTGTGTGT 2475
2492 ATTGCTCTTGTAAACCTCGAGCAAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2551
2476 TTCTTTAAATGATGT 2535
2552 TTCTTCAATATGATGACGCTTAAAGTAACTTCAACCAACATCTGTGATGAGTATGT 2611
2536 CATAAAGCATCTCTAGAGGTGACACAGACTGTAAACCGCATCTGTCCACGCTCTTC 2595
2612 CATAAAGTATATCACAACGTTGACAGGCTCAATCAACGCGCATCTGTGTGTGTGTGTGT 2671
2596 TAGACAAAGCATGAAAGCGGTTTAAACCATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2655
2672 TACGAGGCAAGAGT 2731
2656 GAGACCAAAAGCGGCAAAAGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2715

Db 2732 GGAAGACCAAGCCCAAGCCGAGAGACATCGTTTAACATCTCCGAGGCTGGCAAG 2791
Qy 2716 CAGCTACAGATTGACTACAAAATCAAGAAATCATGACTGGGCTGCATCGCAGAGCTT 2775
Db 2792 CAGCTGAGTTGACTACCGTGAACAGAACTCATGACAGAGACATCTCAGGGCTC 2851
Qy 2776 ACCGGAAGAGCGTTTATGCTGTGAGGTACAAGTCAACAGAAATCACTACTCGAG 2835
Db 2852 ACCCGCAAGGGGATACCGCTGAAGCAGAAAGTGAATGAATCCCTGTATGCCCT 2911
Qy 2836 ACTTGTGAGACGTGAACGTGTACTTACAGCAACAGAAAACGATTTCTGGAAGCG 2895
Db 2912 GCGCTGAGACGTGAATGTACTGTCTGACGCGCACTGAGATAGGCTGTGTGAAAAACG 2971
Qy 2896 CTAGCTGTATCCCTGATTAAGACACTTACACTTAATATCCCGGGATTTTCAAGCT 2955
Db 2972 CTGGCCGCGATCCCTGATTAAGTCTTATCAAAATTCACAGGGTAACTTTACGACC 3031
Qy 2956 TCATTGAGACGACTGCGACGCGAACGACGCCATTATGCGACCGCTTCTGATAGCG 3015
Db 3032 ACATTGAGAAATGCGAAGAAACAGCAAAATTAATGAAGTATTTGAAGACCGCT 3091
Qy 3016 CAGACAGCTGATGTGTTCAGAAATPAGGTGAACGTCTGCTGGCGAAGCTTTAAGCCA 3075
Db 3092 GCGCTGTGACGCGCTTCAGAAACAAAGCGAACGTGTGTGGCGAAAGCCTGTGCT 3151
Qy 3076 GTCTGGCCACGGCCAACTGTGTCTGACGAGACAGAGCTGGAGAC---GTGACACCA 3132
Db 3152 GTCTGAGACCTGCGGATGATGACAGACAGAGAGTGGACACATTAATTAACAGCA 3211
Qy 3133 TTCAAGCATGACAGAGGCTACTCACTGAAATGCGACTGAACCTTTTGGACCAAGTTC 3192
Db 3212 TTTAAGGAGGACAGAGTTACTCTCCAGTGGTGGCTTGAATGAATTTGACCAAGTAC 3271
Qy 3193 TTTGAGTGAAGCTGGAAGTGGTATTTTCCGCTCTACCGTGGCACTTAATTACAG 3252
Db 3272 TATGAGATTGACCTGGAAGTGGCTGTTTCTGCCGGAAGTGTCCCTGATTAACAG 3331
Qy 3253 GATAGACCTGGGATTAACCTGCCAGGGAAGAACTGTATGGGCTTAATAGAGAGTGA 3312
Db 3332 AACAAACACTGGGATTAACAGACTGTGTGAAGATGTATGATTCMAATTCGCAAC 3391
Qy 3313 AAGAGTGTGACCGGATATCCGTGATCAAAAAGCGGTGACACAGCAGGAGTACT 3372
Db 3392 GCCAGCTGGAAGCTAGACATCTTCTGAAAGGGCAGTGGCATCCGGCAGACAGCA 3451
Qy 3373 GATATTAAGAAATTAACCATCAAGGACTACTCTCCAAATTAATGTGTTCATTAAT 3432
Db 3452 GTATTCGAGAAAATAATCCAAACCGCTTCTGTGCTGACAAATGTAAATTCATAC 3511
Qy 3433 CGCCGTTGCCCACTGTTATGTGTGACCAAAAGACAGGATCAACTGATCACAGC 3492
Db 3512 CGCAGGCTGCGACCGCTGTGTGTGATGAAGACGTTAAAGCAGTAGAGGTGAG 3571
Qy 3493 GGATTTCTATTAAGATGAAGGCAAAATGTGTGTGATCGCGATCTTATCAGACT 3552
Db 3572 TGGGTGTGATTAAGTAAAGAGGTACACGTCCTGTGTGATGATTAACCTGGCT 3631
Qy 3553 CCAGGAGAAAGTAAAGTCCATGGGTCAATGGCCCACTAAC-----CATCAG 3603
Db 3632 TTGGCTGACGACAGGGTCACTGGTGTGCACCGCTGAATGCACAGGCGCGATAGTGC 3691
Qy 3604 TGTGATCTCGATTTGGGATACCTAGACGATCGGTAAATATGACATTAATTTGTCAAT 3663
Db 3692 TACACCTTAAGTTAGACTCCGCTGACCGCCGACAGTTGCACTGTGTTGTGAAAC 3751
Qy 3664 GTTAGAGCCCGGTACAGAAACATCACTAACCAAGTGCAGAGATCAGCTATCCACAC 3723
Db 3752 ATTCAACAGGAATTCAGAAATCCACATCAAGCAGTGTGTGACACGCGCATGAAGTG 3811
Qy 3724 AGCATGCTAAAGTAAAGGCTGTCCACCACTGAACACTGCGGAAACATGTGTGCTATA 3783

Db 3812 CAGATGCTTGGGGAGATGCGCTACGACTGTAAAAACCCGCGCATCT---TGATGAGA 3868
Qy 3784 GGGATGCGGCTTGCTGATTCGCGCAACCGAGATATATCATCTGGGGTGGCAGCGCTATT 3843
Db 3869 GCTTACGGATATCGCTTAATAATCAGGAAAGCGCTTGTCTCTTAAAGCAAAAGTTC 3928
Qy 3844 AAGTTTACCCGTGTCTGTCAAGCTTAAGAACACTGCGGAAAAATCTGAGGTTCTTGTG 3903
Db 3929 TCGTGTCAAGAGTGTGGCCCGGAAATGTGTCAACGAAATACAGAAAGTTCCTTCTG 3988
Qy 3904 TTCTTGGGAGAGCAACGCGCAACACACACATGACCGAGACAGACTCGGTATAGTCTT 3963
Db 3989 TTCTTCAACTTTTACCAACGAAAGAGACCTCTACCGTACCAAGATTAACCAACTG 4048
Qy 3964 GACAACATCTATCAAGGGTCAACAGGTACAGAGGAGAGAGCTCCAGGATCAGAGT 4023
Db 4049 AGTCCGTGTATGCGGAGAAAGCATGACACAGCGCGGGTGTGACCATCTACAGATT 4108
Qy 4024 ATCAGAGTGAATTAAGCAAGCGCTGACCAAGCTATCGTTAATGCTTAATAGCAAA 4083
Db 4109 AAGAGACAGACATAGCCACGTGCAAGAGCGGCTGTGTTAAGCAGCTTAAGCGCCGT 4168
Qy 4084 GGTCAACAGGTTCCGAGTGTGCGGTGACCTGTACCGAAATATGCGCGCTCTTTGAT 4143
Db 4169 GGAATGTAGGGAGATGCGGTATCAGAGGCGGTGCGCAAGAAATGCGCGTCAAGCTTTAAG 4228
Qy 4144 AGACAGCAATATGCTGCGGAGCGGTAGACTGTGAAGCAGCAACCG---CTCATCAT 4200
Db 4229 GAGACGAAACACAGTGGGCAATTAATAACGTCAATGTGCGGTGTATCCCGTATC 4288
Qy 4201 CATGCTGAGGACCCAAATTTTCTTAAGATCCGGAACCGGAGGCGCACTTAAGCTGCA 4260
Db 4289 CAGCTGTACGCTTAATTTCTGTGCAAGCTGAAGCGGAGGCGCCGAAATTTGGCC 4348
Qy 4261 GCTGCTCAATGAGACTAGGCTCCATGTCAACGCTGACGGGATTAATAAATATCATGA 4320
Db 4349 GCTGTCAACGGGAGAGGCGCGCGGAATGAACAGACTGTCACTGACACGTAAGCATC 4408
Qy 4321 CCGCTATCTCAACCGGCATCTATTTCTGTGGGCAAAAGATGAGTATCATCTTGCAT 4380
Db 4409 CCGCTGCTGTCCACAGAGTGTTCAGGCGGGAAGATAGCTGACGCAATCCCTCAAC 4468
Qy 4381 CACCTGTTCAGTGTTCGACACTACGAGATCCGATGTCACATATATGCTTGGATAAA 4440
Db 4469 CATCTATTAACAGAAATGAGCGCACGAGCGTGAACGTACATCTACTCAGAGACAAA 4528
Qy 4441 CAATGGAGACAGATATGAGAGCAATTCACCGCAAAAGAAAGCTGCAAAATTTCTGAT 4500
Db 4529 AGTTGGAGAGAAATAATCCAGGAAGCCATTAATGATGAGAGCGGTGTGAGTGTCAAT 4588
Qy 4501 GATGACAAAGCGAGTATGATGATGATGAGGGTCCACCMAACAGCTCTTTGGCAGGC 4560
Db 4589 GATGACTGAGCTGACACAGACTTGTGTAGAGTGAACCCGAGACAGCTGTGTGGGT 4648
Qy 4561 AGACAGGTTTACTCCGTCATATGAGAGGCAAGTTGTATCATACCTGAAAGTACAGATT 4620
Db 4649 CGTAAGGCTATCAGTATCACTGACGCGGTGCTGTACTGTAAAGGATCAAAATTC 4708
Qy 4621 CATGACCGCCAGAGACTTGTCCGAATTCATGACATATGTGGCCCAAAATCTGAGGCT 4680
Db 4709 AACCAAGCTCTATGATTAATGAGAGATGACTGACGTTGTGGCCCAACTGCAAGAGCA 4768
Qy 4681 AATGAGAGATTTGCTGTATCATCTGTGGGAGAGTATGTTCAGACATCCGCTCCAATTC 4740
Db 4769 AACGAAACAGATATGCTATACGCGCTGTGGGAAACAAATGACAAACATCAATCCAATGT 4828
Qy 4741 CCAATAGAGAGTCAAGAGGCTGTCTCACTTCAACACTTCCATGCTGTGTAAATAC 4800
Db 4829 CCGGTGAACGATTCGATTCATCAACACTCCAGACAGTGGCCCTGTCGTGCGCTAC 4888
Qy 4801 GCTATGACGGCTGAGCGCGTATACAGTGTGCGCTGTCCGAAGAAAGCAAGTTGCGCGTA 4860
Db 4889 GCAATGACAGAGAAAGATGCGCCGCTTAGGTCAACACAAAGTTAAAGCATGTGTGTT 4948

1105 AACCAAGGATGTCGTCATATGTAGAGCGCAAGAAATACTTAACAAATGACAGATCT 1164
1165 CTATTACCAAGTGTGCGCCAGGCGTTTTCAGGTGGCGCGTGAAATCTGTCGCACTTG 1224
3782 CTGCTTCGATGTGTGGCGGTGCGATTTAGCAATGTGGGAGGAATCAAGGACGACTT 3841
1225 GACGACGAGAAAGAACTAGGGGTGCGGAGCGCACTTACTATTAGGCTGCTGCGGCT 1284
3842 GATATATAAAAACCTGTGGGTGTCCGAGAGGTCATCTTCTGCTGTGTGGGCA 3901
1285 TTCAAGACCCGAAATACATTCATCAAGAAAGCTGTGTAAGCAAACTTAAGAA 1344
3902 TTTAAAAAGAGAAAGATGACACATGTACAGAAACAGACACCCAGACAAATAGTAG 3961
1345 GTACCTGCGCTTGTGAATCTATTTGTGATTCACGCGCTTACGAGCCAGCGGCTGATATG 1404
3962 GTGCTTCAGAGTTTAATCTGTTGTCATCCGAGCGTATGTTCTACAGGCTCGCAATC 4021
1405 GCGTTCGCGCTAGGCTCAAGCTGTGTTAACCACTGTCAACCCGACCGGCTATT 1464
4022 CCAGTCAGATCACCCATTTAAGATGCTTTGGCCAAAGAACCAAGC--GAGAGTTAATA 4078
1465 ACAATGGCCGATGAGGATCTGCGTGGCTTACAGAGAAAGCTGAAGAGTGGCTGCA 1524
4079 CTGTTCTCGAGCGGTGCTCAACCCAGAGGATGCTGACAAAGAGAGAAAGAGTTGAG 4138
1525 GCGGAGAGATCAGAGAAAGCCCTGCAACCTTGTCTCCCTGAATAAGAAAGAACCGTA 1584
4139 GCGAGCTGACTAAGAAAGCTTACCAACCTCGTCCCATCGCGCGGAGAGCGGA 4198
1585 GAGCAGAAAGTACCTCATTAATGCAAG--GAGAGAGAGTATGAGTGAAGACA 1638
4199 GTCTCGACGTGAGTTGAAGAACTAAGATATCAAGCAGGTGAGGGGTGTGAAACA 4258
1639 CCAGAGGACATCAGGGGTGACAAAGTTACCCAGGCGAAGAAAGATTGGTCTTACGCT 1698
4259 CTGCGACGCGGTGAAGTACCGGACAGCCGAAACGATCTACTAGAAATTTACGTA 4318
1699 ATACTTTCAACCCAGGCGGTATGAATAGTAAAACTGGCGTATCTCACCACTTGGCG 1758
4319 GTTCTGTCCCGGACCGGTCTCAAGAGTCCAAAGTTGGCCCCGCTGACCTCTAGCA 4378
1759 GAACAAAGTATGTAATGACTTCACAAAGTGGGAGGAGATACAAAGTGAAGCCATAC 1818
4379 GAGCAGGTGAATAATATACATTAACGAGAGGCGCGGCTTACCAAGTGAAGGATAT 4438
1819 CACGTAAGTCAATGTATCAGAAAGGAGCGGCGCTCTTTCAAGACTTCAGAGCATTG 1878
4439 GACGCGAGGTCCTACTACATGTGATCGGCAATTCGAGTCCCTGAGTTCAAGCTTTG 4498
1879 AGTGAAGCGCTAGCATGTTTCAACGAGAGGAGTTCGTAAACAGATCTTCACAC 1938
4499 AGCAGAGCGCCATGATGTGTACACGAAAGGAGTTGCTCAAGAGAACTATACAT 4558
1939 ATGCAATCAACGAGAGCGCTAAACCTGACGAGAGTACTATAGACTGTAAAGACT 1998
4559 ATTCCGTTCAACGAGCGCTCTCTGACACGAGCGAGAACTACGAAAGGCTAGAGCT 4618
1999 CAGACACAGATCAGAAATAGTCTGTGATTTGACGAGAAAGTGTTAAGCGGAA 2058
4619 GAAGAAATGTAGCGCGATAGCTGTTCGACGTAGTAAATAATGCTGTGACAGAGAG 4678
2059 GACGAGGTCCTTGTGCTTAACCGGTATCTGTAGATCCACATTTCAAGAGTTGCG 2118
4679 GAAGCGTGGGTTTGTGTGTGTGGAGAGCTAACCAACCCCGTTCCATGAATTTGCC 4738
2119 TAGAGAGTCTCAAGACGACGACGAGCACTTCAAAAGTCCCAACCATGAGTCTAT 2178
4739 TACGAAGGCTGAAGATCAGGCGCGTCCGACCATATTAAGACTAGTAGAGTCTTT 4798

2179 GAGTGCCAGGTTCAAGTAATCTGAAATCATCAAAAGCCGTGTGACTAAGAAAGATCTG 2238
4799 GGGGTTCCGGGATCAGGCAAGCTGTGATTTAAGAGCTCGTACCAAAACGATCTG 4858
2239 GTTGTAGTCCGAAAGAAAGAACTCGCGCAAAATCATAGGATGTAGGAGATGACA 2298
4859 GTCAACGCGGCAAGAGGAACTGCAAGAAATAGTTAACACGTAAGAAAGCAACCGC 4918
2299 CGTATGATTTGCTGTAGACTGTGATTCAGTGTGCTTAATATGGGGTTAAGACCCC 2358
4919 GGGAGGGGACAAAGTAGGAAACAGTACTTCATCTGCTTAACGGGTGCTGTGCGC 4978
2359 GTTAACATCTGTACATTTAGTGAAGCATTTGCTCCATGACAGGAGCGTGTGCGACTG 2418
4979 GTGACATCTCTATATGTGACAGAGGCTTGTGCTTGCATTCGGTACTGTGCGCCCTA 5038
2419 ATTGCACTGTCAAACT--AAGAAAGTGTATTTGTGCGGGGACCCAAACAATGCGC 2475
5039 ATTGCTCTGTAAACCTCGAGCAAAAGTGTGTATGCGGAGACCCCAAGCAATGCGGA 5098
2476 TTCTTTAATGATGTGCTGAAAGTACATTTTAACCATGACATPAGTCAATGAAGTAC 2535
5099 TTTCTCAATATGATGACGCTTAAAGTACTTAACCAACAATCTGCACTGAAGTATGT 5158
2536 CATTAAGCATCTCTAGAGGTGACACAGACTGTAAACCCGATCGTCTCACGCTCTTC 2595
5159 CATTAAGTATATTCAGACGTTTACGCGCTTCAGCTACCGGCTCATCGTGTAGGTTGAC 5218
2596 TACGACAAAGCAATGAAGAGGTTTAACTCATGCTGTATTAATCATATGATACCA 2655
5219 TACGAGGCAAGATGCGCACGACCAACCCGTGCAACAAACCAATATCATATGACACCA 5278
2656 GGGACCAAAAGCGGCAAAAGATGATCTGATTTCTAACTGTTTACAGAGTGGTGAAA 2715
5279 GAGACAGCAAGGCCAAGGACGAGACATGCTTTACATGCTTCGAGGCTGGGCAAG 5338
2716 CAGCTACAGATTGACTACAAATAATCAAGAAATGATGACTCGGCTGATCGAAGACTT 2775
5339 CAGCTGAGTTTGTGACTACGTTGACAGAGTATGATGACAGAGCATCTCAGGGCTC 5398
2776 ACGCGAAAGCGTTTATGCTGTGAGTACAAAGTCAAGAGAAATCCACTTACTGCGAG 2835
5399 ACCGCAAGAGGATATGCGCGTAAAGCAGAAAGTGAATAAATCCCTGTATGCCCC 5458
2836 ACTTCTGAGACGGAAGGTTACTTACAGGACGAGAAAGCAATGCTGGAAGAGC 2895
5459 GCGTGGAGCAGTGAATGTACTGTGACGCGCATGAGATAGCTGTGTGGAAGAACG 5518
2896 CTAGCTGTGATCCCTGATTAAGACATTAAGCTAATATATCCGCGGATTTCAAGGCT 2955
5519 CTGCGCGGATCCCTGATTTAAGGTCTTATCAAACTTCCACAGGTTACTTACGGCC 5578
2956 TCATTGAGCAGCTGCGAGCGCGAAACGAGCCATTATGSCAGCGCTTTGTATAGCCG 3015
5579 ACATTGAAATATGCGAAGAAACGACGAAATATATGAAGTATGAAGAACCGGCT 5638
3016 CAGACAGTATGTGTTCAAGATTAAGTGAAGTGTGCTGGGGAAGCTTTAGAGCCA 3075
5639 GCGCTGTGAGCGGTTCCAGAACCAAGCAAGTGTGTGGGGAAGAAACCTGTGTGCT 5698
3076 GTCTTGGCACCGGCAACATTTGTGCTGACAGACAGAGTGGGAGC--GTTGACCCA 3132
5699 GTCTGAGCACTGCGGAAATCAGATGACAGAGAGAGTGGAGCACATTAATTAAGCA 5758
3133 TTCAAGCATACAGAGGCTACTCACTGAATGGCACTGAATCTTTTGTGACCAAGTTTC 3192
5759 TTTAAGAGGACAGAGCTTACTCTCAGTGTGCTTGAATGAATTTTCCACCAAGTAC 5818
3193 TTTGAGTAACTGTGACAGTGGTTATTTTCCGCTCTTACCGTGCATTAACAGG 3252
5819 TATGAGTTTACCTGAGCAATGTGCTGTGCTTTCGCGGAAAGTGTCCCTGTATTAAGAG 5878
3253 GATCAGACTGTGATTAATCTGCAAGGAAAGACATGTATGGCTTAATAGAGGTAGCA 3312

```

Db      5879  AACAACTGAGGATTAACAGACTGTGGAGAGATGATGATCAATCCGCAACAGCT 5938
Qy      3313  AAGAGTTGTACGCGGATATCCGTGCATCAAAAAGCGTTGACACAGGCGGTAGCT 3372
Db      5939  GCCAGCTGGAAAGTAGACATACCTTCTGAAGGGGAGTGGCATACGGGCAAGAGCA 5998
Qy      3373  GATATAGGAATATATACATCAAGAGCTACTCTCCAACTATATGTGGTTCCATTAA 3432
Db      5999  GTTATCGGAGAAAGAAAATCCACCGCTTTCTGTCTGAGACAAATGTAATCTTATCA 6058
Qy      3433  CGCCGTTGCCCTCCTGTTGATCGTTGACCAAAAAGACAGGGTAACTGATCAACG 3492
Db      6059  CGCAGGCTGCCACGCGCTGTGGTGTAGTACAAAGACGGTTAAAGCGATGAGTTGAG 6118
Qy      3493  GSATTCCTATCTAATGAGGCAATGTGTGTGTGATCGCGGATCTTATCAAGCAT 3552
Db      6119  TGGCTGTCAATTAAGTAAAGAGGGTACACGCTCTGTGTGTGATGATGAACCTGGCT 6178
Qy      3553  CCAGGAAAGAAAGTAGAGTCCATGGGTCATTGCCCATTAATAC-----CATCAGG 3603
Db      6179  TTGCTCGACGAGGAGTCACTTGTGTGTCAACCGCTGAATGTCAACAGGCGCATAGTGC 6238
Qy      3604  TGTGATCTGATTTGGGAATACCTAGCCATGTGGTAAATATGATTAATCTTTGTCAAT 3663
Db      6239  TACGACTTAAGTTTAGACTGCGCGGTACGCGGAGGTGATGACTTGTGTTGTAAC 6298
Qy      3664  GTTAGAGACCCGCTACAGAAACATCACTACCAACGTGCGAGATACGCTATCCACAC 3723
Db      6299  ATTCAACAGGAATTCAGAAATCCACATACAGAGTGTGTGACACCGCCATGAAAGCTG 6358
Qy      3724  AGCATGCTAAGCTGTAAGGCTGTCCACCACTGAACACTGGGCGGAACATGTGTGCTATA 3783
Db      6359  CAGATGCTGTGGGGAGATGCGTATACACTGCTAAACCGGCGGCACT--TGATGAGA 6415
Qy      3784  GGGTATGGGCTTGTGATGCGCAACCGAATATATCATCTGCGGTGCAACGCTCATTT 3843
Db      6416  GCTTACGGAATACCGGATTAATACGCGAAGCGGTTGTTCTCTTAAGCAAGAAAGTTC 6475
Qy      3844  AGTTTACCCGTGTCTGTACGCTTAAGAAACATCCCGAAATATGAGGTTCTTTGCTG 3903
Db      6476  TCGTGTGAAGAGTGTGGCCCGATGTGTACACAGCATACAGAAAGTGTCTTGTGCTG 6535
Qy      3904  TTCTTCGGAAGAGCAACGCGCAACACACATGACAGAGAGACTCGGTGTAGTGGCTT 3963
Db      6536  TTCTCAACTTTGACACGGAAGAGAGACCTTACGCTTACACAGATGATACCAAGCTG 6595
Qy      3964  GACCAATCTATCAAGGCTCAACAGGTACAGAGGAGGAGGACTTCAACGCTACAGAGTG 4023
Db      6596  AGTCCGTGTATGCGGAGAAAGCATGACACGCGCGGTGTGACATCTTACAGAGTT 6655
Qy      4024  ATCAGAGGTGACATTAGCAAGCGCTGACCAAGCTATCTGTAATGTGCTAATAGCAAA 4083
Db      6656  AAGAGAGCAACATAGCCATGTCACAGAAAGCGGTGTGTATAGCAGCTAACCCCGT 6715
Qy      4084  GGTCAACAGGTTCCGAGTGTGCGGTGCACTGTACCCAAATGTCCGCTGTTTAT 4143
Db      6716  GGAATGTAGGGGATGCGTATGACAGGCGGTGCGAAGAAATGCGCTTAAAG 6775
Qy      4144  AGACAGCCAATAGCTGTGCGGACGCTGACTTGTGAAGCAGAACCG---CTCATCAT 4200
Db      6776  GGAGACGACAAACAGTGGGCAATTAABACAGTCAATGTGGGCTCGTACCCGCTATC 6835
Qy      4201  CATGCTGTAGACCAATTTTCTTAAGATGCGGAAACGCGAGGCGACCTTAAGTCCGA 4260
Db      6836  CACGCTGTAGCGCTTAATTTCTCTGCAAGTGAAGCGGAGCGGACCGCAATTTGCC 6895
Qy      4261  GCTGCTCATAGAGATAGCGTCAATCTGCAACGCTGAGCGGATTAACAAAATATCAGTA 4320
Db      6896  GCTGTCTACCGGCGAGTGGCGCGCAAGTAAACAGATCTGTCACTAGCAGCGTACCATC 6955
Qy      4321  CCGTACTGTCAACCGGATCTATTTCTGTGCAAAAGATCGAGTATGCAATCATGTGAT 4380

```

```

Db      6956  CCGCTGTGTCCACAGAGTGTTCAGCGCGGAAGAGATAGGCTCAGCAATCCTCAAC 7015
Qy      4381  CACCTGTCACTGCTTTGACACTAGATGCGGATGCGGATGACCATTAATTTGTTGATAA 4440
Db      7016  CATCTATTCAAGCAATGAGACGCCACGAGCGGTGACGTACCTACTACGAGACAA 7075
Qy      4441  CAATGGAGACAGGATATCGAGGCGCATTCACCCCAAGAAAGCGTGAATTTCTGAT 4500
Db      7076  AGTTGGAGAAAGAAATTCAGGAAGCCATTGACATGAGAGACGCGTGTGAGTTGCTCAAT 7135
Qy      4501  GATGACAAAGCAATGATGACATTAAGTGTGTCAGGTCACCCCAACAGCTCTTTGCGAGC 4560
Db      7136  GATGACGTGAGCTGACCAAGACTTGTGTAGAGTGACCCGGAACAGACGCTGTGGGT 7195
Qy      4561  AGACAGGTTACTCCGTAAATGAGGCGCAATGTATCTATCTCTGGAAGATACAGATTTC 4620
Db      7196  GGTAAAGGCTACAGTACCACTGACGCGGTGCTGTACTGTATGAAAGTACGAATTC 7255
Qy      4621  CATCAGACCGGCAAGGACATTGCGGAATCCATGCAATGTGGCCCAACAAATCTGAGCT 4680
Db      7256  AACAGGCTGTATTTGATATGAGCAGAGATGACAGTGTGTGCGCCAGACTGCAAGAGCA 7315
Qy      4681  AATGACAGATTTGCTGTACATCTGGGGAGAGATATGTCAGATCCGCTCCAAATGC 4740
Db      7316  AACGAACAGATATGCTATATGCGCGTGGCGAAACAAATGACCAATCAGATCCAAATGT 7375
Qy      4741  CCAGTAGAGAGTCAAGAGCGCTTGTCTCACTCAACACTTCCATGCTGTGTATATAC 4800
Db      7376  CCGGGAACGATTCGATTCATCAACACTCCAGAGAGTGCCTGTGCTGCGCTAC 7435
Qy      4801  GCTATGACGCTGAGCGGCTATACAGTGTGCGCTGCGAAGAAAGAACATGTCGCGTA 4860
Db      7436  GCAATGACAGAGAGAGATGCGCGCTTATGATGATACCAAGTTAAAGATGTGTT 7495
Qy      4861  TGTTCATCATTTCTGTTGCCGAAGTACAGATCAACAGCGGTGACAGACTACGTGACG 4920
Db      7496  TGTTCATCTTTTCCCTCCGAAATACCATGATGATGGGGTGCAGAAAGTAAAGTGGCG 7555
Qy      4921  AAACAGTCTGTTTTCAGGCGGTGTACACCGCTGTATACCCCAAGAAAGTACCGGAA 4980
Db      7556  AAGTTCTCTGTTGACCCGACGATCTTCAAGTGTAGTGTCCGCGGAAGATATGCCCA 7615

RESULT 12
ABN86687
ID ABN86687 standard; DNA; 11489 BP.
XX
AC ABN86687;
XX
DT 05-NOV-2002 (first entry)
XX
DE Nucleotide sequence of a pSCA1 suicide DNA vector.
XX
KW Major histocompatibility complex; MHC; antigen presenting cell; APC;
XX antigen; cytotoxic; virulence; gene therapy; CD8; vaccine; therapeutic;
XX cancer; viral infection; ds.
XX
OS Synthetic.
XX
PN WO200261113-A2.
XX
PD 08-AUG-2002.
XX
PF 01-FEB-2002; 2002WO-US02598.
XX
PR 01-FEB-2001; 2001US-26534P.
XX
PA (UWTO ) UNIV JOHNS HOPKINS.
XX
MU Wu T, Hung C;
XX
PI
XX
DR WPI; 2002-619261/66.
XX

```


1832 GAGCGAGTGAATAATATACACATACCGGAGGCGCGGTATACAGGTCCAGCCGATAT 1891
1839 CACGGTAAAGTCATTGTATCCAGAAAGGAGCGGCGTCCGTGTCAAGACTTCCAGGACTTG 1878
1882 GACGGCAGGGTCTCTACTACATGTGAGATCGGCATTCCGGTCCGTGAGTTTCAACCTTGG 1951
1879 AGTGAAGCGCTACGATCGTTTTCACAGAGAGGAGTGTGAAGAGATACCTGAGCCAC 1938
1952 AGCGAGAGGCCACTATGTGTATCAACGAAAGGAGTGTCTCAACAGGAAATATATACAT 2011
1939 ATCGCAATCAACGAGAGAGCGCTAAACACTGACGAAGATCTATTAAGCTGTAAAGACT 1998
2012 ATTCGCGTTCAACGAGACCGTCTGTAACACCGAGAGAACTACGAGAAATCAGAGCT 2071
1999 CAGGACAGAGCTCGAATATAGTCTTCCATATTTAGCGACGAAAGTGTATTAAGGAGAA 2058
2072 GAAAGAACTGACGCGGAGTACGTGTTCAGATGAATAAAAATGCTGCTCAAGAGAGAG 2131
2059 GACGAGGTCCCTTGTGCTTAACCGGTATCTGTAGATCCACATTTTCAACAGTTTGGC 2118
2132 GAAAGGTGGGTTGTGTGTGTGGAGAGCTAACCAACCCCGCTTCCATTAATTTCCGC 2191
2119 TACGAGACTCTCAAGACAGACGACGACGACCTCAAAAGTCCAAACCATCGAGCTAT 2178
2192 TACGAAAGGCTGAAATGATCAGCGCTGCGCACCATATTAAGACTAAGTAGTAGAGTCTT 2251
2179 GAGAGTCCAGGTTCAGGTAAATCTGAAATCATCAAAAGCGCTGTACATTAAGAAATCTG 2238
2252 GGGGTTCCGGATACAGGCAAGTCTGTATTTAAGAGCTGTGTAACCAACACATCTG 2311
2239 GTTGTAGTGCAGAAAGGAAATCGCGAGAAATCATCAGGAGTGTAAAGAGATGAGA 2298
2312 GTACACACGCGAAGAGAGAACTGCCAGAAATAGTTTACGAGTGAAGAGACCGC 2371
2289 GGTATGATGTTGCTGTAGAGCTGTCAATTCAGTCTTCAATGAGGTTAAGCACCC 2358
2372 GGGAGGAGGACAAATGAGGAAACAGTACTCATCTGTCTAAACGGGTTGTGTGTGCG 2431
2359 GTTAACTCTGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2418
2432 GTGACATCTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2491
2419 ATTGCATCTGTCAACT 2475
2492 ATTGCTCTTGTAACT 2551
2476 TTCTTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2535
2552 TTCTTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2611
2536 CATTAAGATCTCTAGAGAGTGCACACAGACTGTAAACCGCATGTCTCCACGCTTTC 2595
2612 CATTAAGATCTCTAGAGAGTGCACACAGACTGTAAACCGCATGTGTCTCTCTCTCTCT 2671
2596 TACGACAAAGCAATGAAGCGTTAAACCATGATGATGATGATGATGATGATGATGATGATG 2655
2672 TACGAGGCAAGATGCGACGACCAACCGTGCACCAACCATATATCATATGACACCA 2731
2656 GGGACACAAAGCCGACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2715
2732 GACAGACCAAGCCGACGAGAGACATCGTGTAAATGATGATGATGATGATGATGATGATG 2791
2716 CAGCTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2775
2792 CAGCTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2851
2776 AGCGGAAAGGCGTTATGCTGTCTGAGTAAAGTCAAGAGATCCACTACTCGAG 2835
2852 ACCCGAAAGGCGTTATGCTGTCTGAGTAAAGTCAAGAGATCCACTACTCGAG 2911
2836 ACTTGTGACAGTGAAGTGTATCTTACGACGACGAAAGGATGCTGTGAGAGC 2895

2912 GCGTCGAGCAGTGAATGATCTGTGACGCGCACTGAGATAGCGTGTGAGAAAAG 2971
2896 CTAGCTGTGATTCCTGATGAAGCACTTACAGTAAATATCCGGGGATTTCCAGGCT 2955
2972 CTGGCCGCGATTCCTGATGAAGTCTTATCAACATTTCCAGAGGATCTTACGGCC 3031
2956 TCATTGAGCAGTGTGACGCGGAAACGACGCCATTATGAGCAGCGCTTGTATGAAGCG 3015
3032 ACATTGAGAGATGAGCAAGAAAGAACAGCAAAATTAAGAGTATGAAGAGACCGCT 3091
3016 CAGACACTGATGTGTCTCAAGATTAAGTAACTGTCTGTGGCGAAGCTTTAGAGCA 3075
3092 GCGCTGTGAGCAGGCTTCAAGAACAAACGATGTGTGGCGAAAAGCCTGTGCT 3151
3076 GTCTTGGCAGCGCAATGTGTGTGACGAGACAGAGTGGAGAC---GTGACACCA 3132
3152 GTCTGACACTGCGGATCAGATTGACACAGAGAGTGAAGCATAATTAACAGCA 3211
3133 TTCAAGCATGACAGAGCTACTACCTGAAATGGCACTGAACCTTCTTTGACACAGTTC 3192
3212 TTTAAGAGAGACAGAGCTTACTCTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3271
3193 TTTGAGTGAACCTGAGCAGTGGTATTTTCCGCTCTTACCGTGCACCTTATTAAG 3252
3272 TATGAGTGTGACCTGAGCAGTGGCTGTCTTCTGCGCGAAGTGTCTCTGTATTAAG 3331
3253 GATCAGCAGTGGATTAATCCGACAGGAAAGCAATGTATGGCTTAATAGAGAGTACCA 3312
3332 AACACCACTGGATTAACAGACTGTGTGAAGAGATGATTAATTAAGCCGCAACACT 3391
3313 AAGAGTGTGACGCGCATCTCCGTCATCAAAAGCGTTGACACAGGAGGTAGCT 3372
3392 GCCAGCTGGAAGTGAACATTAATCTTCTGAAAGGCGAGTGGACATCCGCGAAGAGCA 3451
3373 GATATGAAGAAATTAATCATTAAGAGTACTCTCCCAATTAATGTGTTCCATTAAT 3432
3452 GTTATCGGAGAAAGAAATCAACCGCTTCTGTGTCGACAAATGATTAATCTATCAAC 3511
3433 CGCGGTGCCCCCTCTGTATGTTGATGATGATGATGATGATGATGATGATGATGATGATG 3492
3512 GCGAGGCTGCGCACGCGCTGTGTGTGATGAAGAGCGTTAAAGCAGTAAAGGTTGAG 3571
3493 GGAATCTCATTAAGATGAAGGCAATCTGTGTGTGATGATGATGATGATGATGATGATG 3552
3572 TGCTGTGATTAAGTAAAGAGGATCAACGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 3631
3553 CCAGGAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3603
3632 TTGCTCGACGAGGATCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3691
3604 TGTGATCTGATTTTGGGATTAACCTAGCCATGTGGTAAATATGACATTAATCTTTGTCAAT 3663
3692 TACACCTTAAGTGAAGCTGCGGATGACCGCGAGGTGTGACTGTCTTGTGTGAAC 3751
3664 GTTGAAGACCGGTACAGGAACATCAACCAAGTGTGAGATGATGATGATGATGATGATGATG 3723
3752 ATTACACGGAATTCAGATTCACCACTACAGCAGTGTGTGACACAGCATGAAGCTG 3811
3724 AGCATGCTTAAGTGAAGT 3783
3812 CAGATGCTTGGGGAGATGCGCTAGACTGTGTAAACCCGCGGAGATCT---TGATGAGA 3868
3784 GGGTATGGGCTTGT 3843
3869 GCTTAACGATTAAGCGGATTAAGTGAAGCAGGATGATGATGATGATGATGATGATGATG 3928
3844 AGGTTTAACTGCTGT 3903
3929 TGTGTGTGAAGATGTTGCGCGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3988
3904 TTCTTGGCAGGAAAGCAGGACCAACACATGACAGGACAGACTCGGTGTGTGTGT 3963
3989 TTCTTCAACTTTGACAAAGGAAAGACCTCTAGCGTACACGATGAATACCAAGCTG 4048

QY 3964 GACAACTTATCAAGGGTCAACAGGTACAGGAGGAGGCTCCAGGTACAGATG 4023
DB 4049 AGTCCCTGTATGCCGGAGAACCATGACACGGCGGGTGTGACCACTCTACAGAGTT 4108
QY 4024 ATCAGAGGTGACATTAGCAGAGCGCTGACCCAGCTATCTGTTAACTGTCTATATAGCAA 4083
DB 4109 AAGAGAGAGACATATAGCAGCTGACAGAACGGCTGTGGTTAAGGAGCTTAAGCCCGT 4168
QY 4084 GGTACACCAAGTCCGAGTGTGGGTGCACTGTACCAAAATGCGCGCTTTTGTAT 4143
DB 4169 GGAACGTATAGGGGATGCGGTATGCAAGGCGCTGCGAAGAAATGCGCTTAAAG 4228
QY 4144 AGACAGCAATATAGCTGTGGAGCGGTATGCTTGAAGACGAAACG--CTCATATA 4200
DB 4229 GGAGACAGACACACATGAGGACCAATTAACAGTCACTGCGGTCTGATCCCGTCAATC 4288
QY 4201 CATGCTGTAGGACCAATTTTCTAAGATGCGGAAACCGAGGGCGACCTTAAGCTGCA 4260
DB 4289 CAGCTGTAGCGCTAATTTCTGTGCCACGATGAGCGGAAGGGGACCGGAAATTGGCC 4348
QY 4261 GCTGCTTACATGACATAGCTCCATGCTCAACGCTGAGCGGATTAACAAATATCACTA 4320
DB 4349 GCTGTCTACCGGGAGTGGCCGCAAGTAAACAGACTGTCACTGAGCAGCGTACATC 4408
QY 4321 CCGCTACTGTCAACCGGATCTATTTGTGGTGGCAAAAGATGAGTGAATCATCATTTGAT 4380
DB 4409 CCGCTGCTGTCAACGAGTGTTCAGCCGCGAAGAGATAGGCTGACACCAATCCCTCAAC 4468
QY 4381 CACCTGTCTACTGCTTTCGACATACGAGTCCGATGTCACCATATATTTGCTGGATMA 4440
DB 4469 CATCTATTCAAGACATATGAGACGCAACGCTGACGTCATCTACTGACAGACAAA 4528
QY 4441 CAATGGAGACGAGATATGAGGCCATTCAACGCAAGAAAGCTGAAATTTGTGAT 4500
DB 4529 AGTTGGAGAGAAATCCAGAAAGCCATGACATGAGAGCGGCTGTGAGTGTGTCAT 4588
QY 4501 GATGACAAAGCAGTATGATGATGCTGGTCCAGGGTCCACCAACAGCTTTTGGCAGGC 4560
DB 4589 GATGACGTGAGCTGACACAGACTTGTGTGAGATGACCCGAGACAGACCTGTGGGT 4648
QY 4561 AGACAGGTTACTCCGTCATATGAGGAGCAAGTTGATTCATCTGTAAGTACATCATC 4620
DB 4649 CGTAAAGGCTACATGACACTGAGGGGTGCTGTATCTGTATTTGAAGGTATCGAAATTC 4708
QY 4621 CATGACAGCCCAAGACATTTGCCGAATTCATGCAATGTGGCCCAACAAATCTGAGGCT 4680
DB 4709 AACCAAGCTGCTATGATATGACAGAGATGACTGACGTTGTGGCCCAAGCTCAAGAGCA 4768
QY 4681 AATGACAGATTTGCTGTATCATCTGGGGAGAGATGTCCAGCATCGCTCCAAATGC 4740
DB 4769 AACCAACGATATGCTTATGACCGCTGGCGAACAATGACACATGATTCAAATGT 4828
QY 4741 CCAGTAAAGAGTACAGAGCGTGTGCTCCACTCACAACACTTCATGCTGTGTATATAC 4800
DB 4829 CCGGTGAACGATTCGATTCATCAACACTCCACAGGACAGTGCCTGTGTGGCGCTAC 4888
QY 4801 GCTATAGCGGTGAGCGCGTATACAGTTGGCTGTGCGAAAGAAAGAACTGTGCCGTA 4860
DB 4889 GCAATGACAGACAGACGATGCGCCGCTTAGTTCACACCAAGTTAAAGATGTGTGTT 4948
QY 4861 TGGCTCATATTCCTGTTCCGAGTACAGATCAACAGCGGTGCGAAGACTCAAGTGACGC 4920
DB 4949 TGTCTATTTTTCCTCCGTAATACATGATGATGGGGTGCAGAAAGTAAAGTGCAG 5008
QY 4921 AAACAGTCTGTTTTCAGAGGCTGTACACCGCTGTACACCCAGGAAAGTACCGGAA 4980
DB 5009 AAGGTTCTCTGTTGACCCGACGCTATCTTCAAGTGTAAATGCGCGGAAGTATGCCGA 5068

RESULT 13
ABN86690
ID ABN86690 standard; DNA; 13599 BP.

XX AC ABN86690;
XX DT 05-NOV-2002 (first entry)
XX DE Nucleotide sequence of vector pSCA1-E7-Hsp70.
XX KM Major histocompatibility complex; MHC; antigen presenting cell; APC;
XX KM antigen; cytosolic; vinidic; gene therapy; CD8; vaccine; therapeutic;
XX KM cancer; viral infection; HPV; E7; heat shock protein 70; Hsp70; de.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT misc_feature 7405..9509
XX FT /*tag= a
XX FT /note= "E7-Hsp70 fusion sequence"
XX PN MO20261113-A2.
XX PD 08-AUG-2002.
XX PE 01-FEB-2002; 2002WO-US02598.
XX PR 01-FEB-2001; 2001US-265334P.
XX PA (UYJO) UNIV JOHNS HOPKINS.
XX PT Wu T, Hung C;
XX DR WPI; 2002-619261/66.
XX PT Nucleic acid molecule encoding a fusion polypeptide that promotes
PT processing via the Major Histocompatibility Complex class I pathway
PT and/or promotes activity of an antigen presenting cell, useful as
PT vaccine for cancer and viral infections -
XX PS Claim 37; Page 34-37; 127pp; English.
XX CC The invention relates to a new nucleic acid molecule (I) encoding a first
CC fusion polypeptide useful as a vaccine composition. (I) comprises a first
CC nucleic acid sequence encoding a first polypeptide or peptide that
CC promotes processing via the Major Histocompatibility Complex (MHC) class
CC I pathway (MHC-I-pp) and/or promotes development or activity of an
CC antigen presenting cell (APC). The second nucleic acid sequence is linked
CC in frame to the first nucleic acid sequence or to a linker nucleic acid
CC sequence and encodes an antigenic polypeptide or peptide. The methods and
CC compositions of the present invention are useful as therapeutic vaccine
CC for cancer and for major viral infections, such as hepatitis and cervical
CC cancer, that cause morbidity and mortality. They can also be used in
CC treating animal diseases, such as equine herpesvirus, bovine viruses,
CC Marek's disease, retroviral and lentiviral diseases and rabies, in the
CC veterinary medicine context. The present sequence represents the
CC nucleotide sequence of vector pSCA1-E7-Hsp70 comprising the human
CC papillomavirus (HPV) E7 antigenic protein and M. tuberculosis heat
CC shock protein 70 (Hsp70) fusion sequence.
XX SQ Sequence 13599 BP; 3582 A; 3516 C; 3647 G; 2854 T; 0 other;
Query Match 31.2%; Score 1562.4; DB 24; Length 13599;
Best Local Similarity 58.3%; Pred. No. 0;
Matches 2903; Conservative 0; Mismatches 2041; Indels 36; Gaps 8;
QY 31 AGAATTACGCTTGAATGATGCTGACAGCCCGTATGTCAGTGTTCACGCGGAGCTTT 90
DB 95 AAATGATGATGTTGATTTGAGGCTGACAGCCCATTCATCAAGTCTTTCAGAAAGCAATTT 154
QY 91 CCACAAATTTGAGATGGAAGAGGAGGAGTCACTGACATGACATGCAATGCGAGAGCG 150
DB 155 CCGTGTTCAGAGTGTGATCTTTCAGATGACACCAATGATGCAATGATGCAATGCAATG 214
QY 151 TTTTGCATGTGGCAACAAAGCTCATTTGAGAGCGCAAGTGCACCGGACCAAGTATCTTG 210

D 215 TTTTCGACCTGGCTACCAAAATTGATCGAGCAGAGACTGACAAAGACACTCATCTTG 274
Q 211 GACATTGGAAGTGGCCCGCTCAGACATGACATTCCAATCACCAGTATCATTTGATCGC 270
D 275 GATATCGCAGTGGCCCTTCCAGAGAAATGATGTCTACGACAAATACCACTGCGTATGC 334
Q 271 CCTATGATTAAGCGCTGAAGACCAGACATCAACCGTATGCAAGAAAGCT-----T 324
D 335 CCTATGCGCAGCGCAGAAAGACCCCGAAAGGCTCGATAGTACGCAAAAGAACTGGCAGCG 394
Q 325 AAGAAAATGACATTACGACAAAGAACTAGCCTCTAAGCGCGCAGACTGCTGGAATGC 384
D 395 GCCTCCGGAGAGTCTGATAGAGATCGCAGAAAATCAACCGACCTGCAGACCGCT 454
Q 385 ATGTCAACCCAGACGACAGACTCCATCTGTGTATGACACAGAGCGCCAGTGTAGG 444
D 455 ATGGCTAGCCAGACGCTGAATCTCTTACCTTTGCTGATACAGACCTCACGCTGCT 514
Q 445 TACTTTGAAAGTGAAGATATACCAAGATGTGTACGAGTCATGCAACGACCAATC 504
D 515 ACGGACCGCAAGTGGCGGTATACCAAGAGCTGTATGCTGTACATGCAACATCGCTG 574
Q 505 TACCAACGAGCGCTTAAGAGATTTAGACATTTACTGATAGCGCTTTGACACGACCT 564
D 575 TACCATCAGGCGATGAAAGGTGTCAAGACGCGGTATTGGATTGGTTGACACACCCG 634
Q 565 TTTATGTACAAAACATGGCAGGTTCTTACCCTTCAACAGCAAGCTGGGCTACGAG 624
D 635 TTTATGTTTGAACGCGCTACGAGCGGCTGTATCCAACTACGCAAACTGGCCGACGAG 694
Q 625 AGAGATTGGAAGACGATTAACATTTGCGCTCGTAACTCAGATCTTCAGAGAGCAGGCT 684
D 695 CAGGTGTAACAGGCGCAGGAACATAGGACTGTGTGAGACATCTTGAAGTGAAGGAAAGCTC 754
Q 685 GGAACACTCTCATCTTATGAGAAAGAGGCTCAACTCTAATTAATGATCATATTCTCG 744
D 755 GCGAAACTGTCTATCTCCGCAAGAGCAATTTGAACCTTGGCAGACAGTCAATGTTCTCG 814
Q 745 GTTGTTCAACAATCTACACAGAAATAGATCATGTTACGTAGTGGCATTTCCAAAC 804
D 815 GTAGATCTTCACTTGTACCTGAGACAGAAAGCTACGAGAGCTGGGACTTACCTTCC 874
Q 805 GTGTTCACTTGAAGAAAGTCTTAACCTTACAGGTAGTGTGGAGCACTTGTACAGTGT 864
D 875 GTATTCAACCTGAAGGTAAACAATCTTACCTGTAGGTGCAATACATCTATCATGT 934
Q 865 GAAGGTACGTCTAATAAAGATTAACATCAGCCCAAGCATATACGTTAAAGTTAGAAC 924
D 935 GAAGGTACGTATGTTAAGAAATCACTATGTGCCCCGCTGTACGTTAAACGCTAGAG 994
Q 925 TTGGGTCACAAATGATCGGAGGGTTCTTGAATGTGAAGTCAACAGATACGCTGGC 984
D 995 TACGCGTAGCATACCGGAGGGATTTCTAGTGTGCAACACAGACACTGTCAA 1054
Q 985 GCGCAGAGGGTTCTTTTGT 1044
D 1055 GGAAGAAAGTCTTATCCCTGTATGCACTACCTCCCTCAACATCTGTGATCAATG 1114
Q 1045 ACAGGATTTCTGGCACTGACGTTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1104
D 1115 ACTGCACTATGACGACGACGCTCACACCGAGAGCGACAGAAAGTTGTAGTGGATTG 1174
Q 1105 AACCAAGATTTGT 1164
D 1175 AATCAGAGATTAAGT 1234
Q 1165 CTATTACAGT 1224
D 1235 CTGCTTCGATTTGT 1294
Q 1225 GACGACGAGAAAGAACTAGGGGTGTGGGAGCGCACTTTACTATGGGCTGTGTGTGTGT 1284
D 1295 GATGATGAAAAAATCTGTGGGTGTCCGAGAGAGTCACTTACTGTGTGTGTGTGTGTGTGT 1354

Q 1285 TTCAAGACCCAGAAAATCACAATCCATCTACAGAAAGCTGTGTAACGCAAACTTAAGAA 1344
D 1355 TTTAAAGAGAGAAATGACACCACTGTACAGAAACCAAGACCCACAGAAATAGTAAG 1414
Q 1345 GTACCTGCGCTTTTGAATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1404
D 1415 GTGCTTCAAGATTTAACTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1474
Q 1405 GGTTCGCGCGTAGGCTCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1464
D 1475 CCAGTCAGATCAAGGATTAAGATGCTTTTGGCCAGAAAGAACCAACG---GAGATTAATA 1531
Q 1465 ACAATGCGCATGTGTGAGCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1524
D 1532 CTGTGTTCTGACGCGT 1591
Q 1525 GCGGAAGATCAGAAAGCCCTGCACTTGTCTCCCTGAATTAAGAAAAGAGCCGTA 1584
D 1592 GCCGAGCTGACTAGAGAAAGCTTACACCCCTGCTGCCATCGCGCGCGGAGAGCGGGA 1651
Q 1585 GAGCAGAAAGTACCTCATTTATGTGAAG-----GACGAGCAGGTAGCTGTGAAGACA 1638
D 1652 GTCTCGACGTGTGAAGAACTAGAGTATCACGAGGTGTGAGGGGTGTGTGAACA 1711
Q 1639 CCAGAGAGACATCAGGGTGTCAAGTTAACCGAGCGAAGAAAGATTGGGTCTTACGCT 1698
D 1712 CTTGCGACGCGTTGAAGTCAACGACGCGCAACGACGTAATCTGTGAATTTACGTA 1771
Q 1699 ATACTTACCCCAAGCGGTATGTAATGTGAATAATCGCGGTGTATCCACCATTTGCG 1758
D 1772 GTTGTCTCCCGCAGACCGTGTCTCAAGACTTCAAGTTGGCCCGCTGTACCTCTTACGA 1831
Q 1759 GACAAGTACTGTATATCTCAAAAGGTGTGGGAGGAGATTAAGTGAAGCATAC 1818
D 1832 GAGCAGGTGAATAATTAATTAACATTAACGAGGCGCGCGGTACAGGTGTGACGGAAT 1891
Q 1819 CACGTAAGTCTATGTATCTCAAGAAAGGAGCGCGGCTCTGTCAAGACTTCAGGATTTG 1878
D 1892 GAGGCAAGGTCTTATCTACATGTGTATCGGCTATTTCCGCTCTCTGAGTTTCAAGCTTTG 1951
Q 1879 AGTGAAGCGGTACGATCGTTTTCACAGAGAGGAGTGTGTAAACAGATACCTGACAC 1938
D 1952 AGCGAGACGCGCATATGT 2011
Q 1939 ATCCGAATCAACGAGAGCGCTTAACCTGACGAGAGTACTATTAAGCTGTAAAGCT 1998
D 2012 ATTGCGTTCAACGAGCGCTGCTGTGAACCGAGGAGAACTACGAGAAAGTCAAGACT 2071
Q 1999 CAGACACAGACTAGAAATAGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2058
D 2072 GAAGAACTGACCGCGATAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2131
Q 2059 GACGAGGTCTCTTGTGCTTAAACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2118
D 2132 GAAGGTGTGGT 2191
Q 2119 TACGAGTCTCAAGACAGACAGACACTTCAAAAGTCCCAACCATGGAGTCTAT 2178
D 2192 TACGAAAGGCTGAAGATCAGCGCTCGGACCAATATAGACTACAGTGTGTGTGTGT 2251
Q 2179 GAGAGTCCAGGTTCAAGTAAATCTGTGAATCAATCAAAAGCGCTGTGTGTGTGTGTGT 2238
D 2252 GGGGTTCCGGATCAGGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2311
Q 2239 GTTGTAGTGTGGAAGAAAGAACTGCGCAGAAATCATCAGGAGTGTGAAGAGATGAGA 2298
D 2312 GTCAACAGCGGCAAGAGAGAACTGTCCAGAAATAGTTAAAGAGCTGAAGAGACCGC 2371
Q 2299 GGTATGATGT 2358
D 2372 GCGAAGGAGCAAGTAGGAAAGAAAGATGATCTCTGTAAACGGGTGTGTGTGTGTGT 2431

QY 2359 GTTAACACTGTGATGATGAGCATTTGCTGCGCATGCGAGGACGCTGCTGACATG 2418
Db 2432 GTGACATCTCATATGATGAGAGGCTTTGCTGACCATTCGGTACTCTGCTGCTT 2491
QY 2419 ATTGCCATCGTCAACCT--AAGAAAGTGATTTGTCGGGGAGCCCAAAACAATGCGGC 2475
Db 2492 ATTGCTCTGTTAAACCTCGAGCAAAAGTGTTATGCGAGACCCCAACCAATGCGGA 2551
QY 2476 TTCTTTAACTGATGCTGCTAAAGTACTTTTAAACCTGACATATGACTGAAGTGTAC 2535
Db 2552 TTCTTCAATATGATGACGCTTAAGGTGAATTCACCAACAACATCTGACGAAAGTATGT 2611
QY 2536 CATAAAGCATCTCTAGAGAGTGACACAGACTGTAAACCGCATGCTTCCAGCTCTTC 2595
Db 2612 CATAAAGTATATCAGACGTTGACCGGTGCACTGACGCGCATCTGTGTCTACGTTGAC 2671
QY 2596 TAGACAAGCGAATGAAGCGTTAAACCATGTGCTGTATMAATCATATGATACCA 2655
Db 2672 TAGGAGGCAAGATGCGCACGACCAACCGTGCAACAAACCATATATCATGACACCA 2731
QY 2656 GGGACACAAAGCCGCAACAAAGATGATGATTTCTAACCTGTTTCAAGGATGGGTAAA 2715
Db 2732 GGACAGACCAAGCCCAAGCCAGAGACATCGTTTAAACATGCTTCCAGGCTGGGCAAG 2791
QY 2716 CAGCTACAGATTGACTACAAAATCAAGAAATCATGACTGCGGCTGATCCGACGACTT 2775
Db 2792 CAGTTCAGATTGACTACCGTGACAGAACTCATGACAGACAGATCTCAGAGGCTC 2851
QY 2776 ACGCGAAAGCGTTTATGCTGTCAAGTCAAGGAAATCCACTCTACTCGAG 2835
Db 2852 ACCCGCAAGGGGTATACGCGTAAAGGCAAGGATGAAATCCCTTGTATGCCCT 2911
QY 2836 ACTTCTGACGCGTGAACGTTACTTACAGCAACAGAAACCATTTGCTGGAAGCG 2895
Db 2912 GCGTGGAGCAGTGAATGTACTGCTGACGCACTGAGATGAGCTGCTGTGAAGAAACG 2971
QY 2896 CTAGCTGCTGATCCCTGGATTAAGACACTTACACTTAATATCCCGGGATTTTACAGGCT 2955
Db 2972 CTGCGCGCGCATCCCTGGATTAAAGTCTTATCAAAACATTTCCACAGGTTACTTTACGGCC 3031
QY 2956 TCATTTGACGACTGCGACGCGGAACAGACGCCATTATGACGCGTTCTTGATTAAGCG 3015
Db 3032 ACATTTGGAAGATGGCAAGAAAGAACAGCAAAATTAATGAAGTGAATGAAGACCGGCT 3091
QY 3016 CAGACAGCTGATGTGTTCCAAATTAAGTGAACGTCTGCTGGGGAAGCTTTAGAGCA 3075
Db 3092 GCGGCTTGAGCGGTTCCAAACAAAGCGAAGCTGTGTGGGAAAGCGCTGTGCT 3151
QY 3076 GTCTTGGCCACGCGCAACATTGTGCTGACGAGACGACGTGGAGAC--GTTGACCCA 3132
Db 3152 GTCTTGACACTGCGCGAATCAGATTGACACAGAGAGTGGACACCATATTAACAGCA 3211
QY 3133 TTCAAGCATGACAGAGCTACTCACTGAAATGCGACTGAACCTTCTTTTGACACAGGTTG 3192
Db 3212 TTTAAGGAGGACAAGCTTACTCTCAAGTGTGGCTTGAATGAATTTGACCAAGTAC 3271
QY 3193 TTTGAGAGAGCTGAGACAGTGGGTATTTTCGCTCTTACCGTGCATCTACTACAG 3252
Db 3272 TATGAGTTGACCTGAGACAGTGGCTGTTTTCTGCGCGAAGGTGCTCTGTATTACAG 3331
QY 3253 GATGAGCACTGGGATTAATCTGCCAGAGGAACAATGATGGGCTTAATAGAGAGTAGCA 3312
Db 3332 AACAACACTGGGATTAACAGACTGTGTGAAGAGATGATGATTCAAAGCCGCAACAGCT 3391
QY 3313 AAGAGTTGTACGCGCATATCTGTCATCAAAAGCGTTGACACAGGAGGATGCT 3372
Db 3392 GCCAGGCTGGAAGCTAGACATACCTTCTGAAGGGGAGTGGACATACGCGCAACAGGCA 3451
QY 3373 GATATTAAGATTAATACATCAAGACTACTCTCAACAAATTAATGTGTTCCATTAAT 3432
Db 3452 GTTATTCGAGAAAGAAATCCAAACCGCTTTCTGTGTGACCAATGTATTAATCTATCAAC 3511
QY 3433 CGCGGTTGCCCACTGTTGATGTTGACCAACAAAGAGGATTAACATGATCACAGC 3492

Db 3512 CGAGGCTGCGCACCGCCTGTGTGCTGATGATACAAAGCGTTAAAGCACTAGGTTAG 3571
QY 3493 GAATCTCTATCTAATGAAAGGCAAAATCTGTGTTGTGATTCGGCATCTTATCAGCATT 3552
Db 3572 TGGCTGTCTAATTAAGTAAGAGGTACCAAGTCTCTGTGTGATGATGATCAACCTGGCT 3611
QY 3553 CAGGGAAGAAAGTAAAGTGCATGGGTCCATTCGCCCATATAC-----CATCAGG 3603
Db 3632 TTGCTTGACGACAGGTCATCTGGTGTCTACCGTGAATGTCAAGCGCCGATAGGTGC 3691
QY 3604 TGTGATCTGATTTTGGGAATACCTAGCCATGTGCTTAATATGACATTAATCTTTGCAAT 3663
Db 3692 TAGACCTTAATTTAGACTGCGCGGTGAGCGCGGAGGTTGCACTTGTGTGTAAC 3751
QY 3664 GTTAGAGACCCCGTACAGAAACATCACTACCAACATGTCGAGATACCGTATCCACAC 3723
Db 3752 ATTCAACGGAATTCAGAAATCCACACTACACAGTGTGCAACACGCAATGAACTG 3811
QY 3724 AGCATGCTAAGTGAAGGCTGTCCACCACTGAAACCTGGCGGAACATGTGGCTATA 3783
Db 3812 CAGATGCTTGGGAGATGCGCTTACACTGTAAACCGGCGGATCT--TGATAGA 3868
QY 3784 GGGTATGGGCTTGTGATTCGCGCAACGAGAAATATCACTGCGGTGCGACGCTCATTT 3843
Db 3869 GTTACGATATCGCGATTAATACAGGAAGCGGTGTTCTTCTTAACAGAAAGTTTC 3928
QY 3844 AGCTTACCCGTGTCTGTACGCTTAAGAACATGCGGAAATATCTGAGTTCCTTGTG 3903
Db 3929 TCGTTCGAAGAGTGTGCCCGGATTTGTGTACCAAGCAATTCAGAAAGTGTCTTGTG 3988
QY 3904 TTCTTGGCAGAGCAACGGCAACACACATGACAGAGACAGTCCGCTGTAGTCTT 3963
Db 3989 TTCTTCACTTTTGAACAGAAAGACCTCTACGCTAACAGATGAATACCAAGCTG 4048
QY 3964 GACAACATCTATCAAGGTTCAACAGGTACAGAGCAGAGAGACTCCAGGTTACAGATG 4023
Db 4049 AGTGCCTGTATGCGGAGAAAGCATGACACAGGCGGGTGTGACCATCTACAGATT 4108
QY 4024 ATCAGAGTGAATTAAGCAAGAGCGCTGACCAAGCTATGCTTAATGCTTAATAGCAA 4083
Db 4109 AAGGACACACATATAGCAGCTGACAGAAAGCGGCTGTGTTAACGAGTTAACGCCGT 4168
QY 4084 GGTCAACGAGTTCCGAGTGTGCGGTGACCTGATCCGAAATAGGCGGCTGCTTTGAT 4143
Db 4169 GGAATCTAAGGATGCGTATGACAGGCGCTGTGCAAGAAATGGCGCTTAAG 4228
QY 4144 AGACAGCAATAGCTGTGCGGACGGCTAGAATTGTGAAGACGAACCG--CTCATATA 4200
Db 4229 GAGACACAAACACAGTGGGACAAATTAACAGTCAATGTGCGGTGTACCCGTCATC 4288
QY 4201 CATGCTGTAGAACCAATTTTTCTAAGATGCGGAAACCGGAGGCGCACTTAAGCTGCA 4260
Db 4289 CACGCTGTACGCTTAATTTCTGTGCAACCTGAAACCGGAAGGAGCCGCAATTTGGCC 4348
QY 4261 GTCGCTAATGACATAGCGTCCATGTCACAGCTGAGGAGTTACAAATAATACGTA 4320
Db 4349 GCTGTCAACGGGACAGTGGCGCGCGGAAGTAAACAGACTGTCACTGACAGCGTAGCCATC 4408
QY 4321 CCGCTACTGTCAACCGGCATCTATTCTGTGTGGAAGAAAGATGAGTATGCAATTTGCAT 4380
Db 4409 CCGCTGTGTCAACAGAGAGTGTCAAGCGGCGGAAGATATGAGCTGACGACATCCCTCAAC 4468
QY 4381 CACTGTTCACCTGTTTGCACATAGCGATGCCGATGTCACTATATTTGCTTGGATAAA 4440
Db 4469 CATCTATTTCAAGCAAGAGCGCACGAGCGCTGACCTGACATCTACAGAGACAAA 4528
QY 4441 CAATGGAGACCGAGATATCGAGCCATTCACCGCAAAAGAGCGTCAAAATTTGAT 4500
Db 4529 AGTTGGAGAAAGAAATTCAGAAAGCCATTTGACATGAGACGCGTGTGGAGTTGCTCAAT 4588
QY 4501 GATGACAGCGAGTAGCATTTGATGAGGTTCAACCAACAGCTTTTGGAGGC 4560

Db 4589 GATGACGTGAGCTGACCAAGACTTGTGAGAGTGACCCCGAGACGACCTGTGGGT 4648
Qy 4561 AGACCAAGTCTACTCCGTAATGAGGCAAGTGTATTCACTTACCTGGAAGGTACACGATT 4620
Db 4649 CGTAGGCTACAGTACCACTGACCGGTGCTGCTACTGTTGAAAGTACGAAATTC 4708
Qy 4621 CATCAGACCGCCAGACGACATTGCCGAATTCATGCAATGTGGCCCAACAAATCTGAGCT 4680
Db 4709 AACCAAGCTGTATTGATATGACAGATATGACGTTGTGCCCCAGACTGCAAGAGCA 4768
Qy 4681 AATGACGATTTGTTGTGACATCTCTGGGGAGAGATATGTCCAGATCCGCTCCAAATGC 4740
Db 4769 AACGACGATATGCTTATACCGCTGCGGCAAAATGAGCAAAATGATCCAAATGT 4828
Qy 4741 CCAGTACGAGAGTACGAGCGCTGCTCCACTCAACACTTCATGCTGTGTAATAC 4800
Db 4829 CCGGGAACGATTCGATTCATCAACCTCCAGAGACATGCCCTGCTGCGCTAC 4888
Qy 4801 GCTATGACGGCTGACGGCGTATACAGTTGCGCTCTGCGAAGAAAGAACAGTTCCCGTA 4860
Db 4889 GCAATGACAGGACGAGATGCGCGCTTAGTACACACAAAGTTAAAGCATGTGTT 4948
Qy 4861 TGCTATCATCTCTGTTGCCGAATGACAGATCAAGGCTGACAGAGCTACAGTGCAC 4920
Db 4949 TGCTATCTTTTCCCTCCGAAATACATGATGAGGTGACAGAGGTAAAGTGCAG 5008
Qy 4921 AACACGTCCTGTTTTCAGGCGGTGTACACCGGCTGTACACCCAGAGAGTACGGGAA 4980
Db 5009 AAGGTTCTCTGTTGACCCGACGATCCTTAGTGTAGTCCCGGAAAGTATGCCGA 5068

RESULT 14
AAQ26021
ID AAQ26021 standard; DNA; 11517 BP.
XX AAQ26021;
AC 25-MAR-2003 (updated)
DT 05-JAN-1993 (first entry)
XX
DE pSP6-SFV4 RNA transcript as DNA.
XX
KW Semliki forest virus; SFV; SP6 expression vector; RNA polymerase; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 87..7379
FT /tag= a
FT /label= Non-structural_polyprotein
FT CDS 7421..11182
FT /tag= b
FT /label= Structural_polyprotein
XX
XX MO9210578-A1.
XX
XX 25-JUN-1992.
XX
XX 12-DEC-1991; 91WO-SE00855.
XX
XX 13-DEC-1990; 90SE-0003978.
XX
XX (BIOP-) BIOPTRON AB.
XX
XX Garoff H, Liljestrom P;
XX
XX WPI; 1992-234633/28.
XX
XX P-PSDB; AAR25138, AAR28337.
XX
XX RNA mol. derived from alphavirus RNA genome - chimeric alphavirus
XX
XX antigen and vaccine for immunisation against viral infections
XX
XX Disclosure; Fig 5; 94pp; English.

XX
CC The sequence given contains a full length Semliki forest virus (SFV)
CC cDNA clone within an SP6 expression vector. The SP6 RNA polymerase
CC promoter allows in vitro transcription of full length and infectious
CC transcripts.
CC (Updated on 25-MAR-2003 to correct PN field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 11517 BP; 3130 A; 2985 C; 3091 G; 2311 T; 0 other;

Query Match 31.1%; Score 1554.4; DB 13; Length 11517;
Best Local Similarity 58.2%; Pred. No. 0;
Matches 2898; Conservative 0; Mismatches 2046; Indels 36; Gaps 8;

Qy 31 AGAATTCAGTGAATTAGATGCTGACGCCCCGATGTCAGTCTGTTACAGCGGCTT 90
Db 96 AAGTGCATGTTGATATGAGGCTGACGCCCCATTCATCAAGTCTTTGCAAGGCAATTT 155
Qy 91 CCACAATTTGATGCAAGCAAGCAGGTCACTGCAATGACATGTCATGCCAGAGCG 150
Db 156 CCGTCTGAGGTGAGTCAATTCAGGTCAACCAATGACATGCAAAATGCCAGAGCA 215
Qy 151 TTTTGCATGTGGCAACAAAGCTCATTGAGAGCGAAGTGCACCGGACCAAGTTATCTTG 210
Db 216 TTTTGCACCTGGCTACCAATTTATGAGCAGAGAGATGCAAAAGCACACTCATCTTG 275
Qy 211 GACATTTGAATGGGCGCGGTGACATGACATTCATCAACCGTATCTGTATCTGC 270
Db 276 GATATCGGCACTGGCGCTTCAGAGAAATGATGTCTGACCAATATCACCTGCTATGC 335
Qy 271 CCTATGATAGCGCTGAAAGACCCGACAGCTACACCGTATGCAAGAAAGT-----T 324
Db 336 CTTATGCGCAGCGAGAAAGACCCGAAAGGCTCGATGCTACGCAAAATGCGCAGCG 395
Qy 325 AAGAAAAGTGAATTACCGACAGAAACATAGCTCTTAAGGCGGACGACCTCTGGAATC 384
Db 396 GCTTCGGGTAGGTGCTGATGAGAGATGCGAGAAATATCACCGACTCTGACACGCTG 455
Qy 385 ATGTCAACACGACGAGACATCTCTGTGTATGACACAGAGCCACGTTAGG 444
Db 456 ATGGCTAGCCAGACGCTGAATCTCTTACCTTTGCTCGATACAGAGTCAAGTGTGTGT 515
Qy 445 TACTTTGGAAGTGTAGAGATATACCAAGATGTGACGAGTCCATGACGACGATCAATC 504
Db 516 ACGGACCGCAAGTGTGCGCTATACAGACGCTGTATGCTATGACCAACATCTGCTG 575
Qy 505 TACCAACAGGCGCTTAAGAGATTAGACAAATTTACTGATAGCTTTGACACGACCCCT 564
Db 576 TACCATCAGGCGATGAATGTGTCAAGACGCGTATGATTGGTTTGACACACCCCG 635
Qy 565 TTTATGTACAAAACATGGCAGGTTCTTACCTTACCAACAGAACTGGGCTGACGAG 624
Db 636 TTTATGTTTACGCGCTAGACGCGCGCTATCAACTACGACCAAACTGGGCGCACGAG 695
Qy 625 AGAGATTGGAAGACGTAATGCTGCGTGAATCTAGATCTTCAGAGAGACGCTT 684
Db 696 CAGGTGTACAGCGCAGAACATAGACTGTGTACAGATCTTGTACTGAGGAAAGATCTC 755
Qy 685 GAAAACCTCTCAATCTTAGAAGAGAGGCTCCACCTTAATTAAGATCATATTCTCG 744
Db 756 GGCAAAATGCTCATTTCTCCGAAAGAAATTAAGAACTTGGCAGACATGATGTTCTCG 815
Qy 745 GTTGTTCAACATCTACAGAAATGATCACTGTTTACGTAGCTGACATCTTCAAAAC 804
Db 816 GTAGATCTACATGTGTACACTGAGAGCAAGAAAGTACTAGAGACTGCACTTACCTTC 875
Qy 805 GTGTTCACTGAAGAAAGTGAATCTTCACTGACAGTGTGAGGACCATGTTACAGCTGT 864
Db 876 GTATTTCACTGAAGAAAGTGAATCTTCTTACCTGTAGGTGCAATCCATGTATCATGT 935
Qy 865 GAAAGGTACGTCATCAAAAGATTAAGATCAGCCAGAGACTATACGTTAAAGTTGAGAAC 924
Db 936 GAAAGGTACGTTAAGAAATCATATGTGCCCCGCGCTGTACGGTAAACGTTAGGG 995

QY 925 TTGCGTCCAAATGCAATCGGAGGGTTTCTTGAGTTGCAAAAGTCACAGATACCGTGGC 984
Db 996 TACCGCGTGAAGTATCAACGCGGAGGATTCCTAGTGTCAAGACACAGACACTGTCAA 1055
QY 985 GCGCAGAGGGTTCTTTTCTGTGTGTACGATGTACAGCCACACTTGTGGATCAATG 1044
Db 1056 GGAAGAAAGTCTATTCCTGTATGTGACCTACCTCCCTCAACCATCTGTGATCAATG 1115
QY 1045 AACAGGATTCGGAATCTGAGTGTAGTGTGATGATGACGACAAAGAACTATGTGGGTC 1104
Db 1116 ACTGGCACTACGAGGACCGAGTCACACCGGAGAGCGACAGAAAGTTGTATGAGGATG 1175
QY 1105 AACCAAGATTTGCTGCTCAATGTAGGACGCAAAAGAAATCTAACACAAATGACAGAT 1164
Db 1176 AATAGAGATGTGTGTGACGGAAGAACACAGGAAACATTAACGATTAAGATCTAT 1235
QY 1165 CTATTAACAGTGTGCGCCAGGCGTTTTCAGGTGGGCGGTGAACATCGTCCGACTTG 1224
Db 1236 CTGCTCCGATTTGTGGCGGTGCAATTTAGCAAGTGGGGAGGGAATACAAAGCAGACTT 1295
QY 1225 GACGACGAGAAAGAACTAGGGGTGCGGAGCGCACTTTACTATGGGTGCTGCTGGCT 1284
Db 1296 GATGATAAAAAAGCTGTGGGTGTCCGAGAGGTCACTTACTTGTCTGTGTGTGCA 1355
QY 1285 TTCAAGACCCAGAAATCAGATCCATCTACAAAGAACCTGTAGCAAAATTAAGAA 1344
Db 1356 TTTAAAGAGAGAAAGATGACACCAATGTACAAAGAACAGACACCAACATATGTAG 1415
QY 1345 GTACCTGCGCTCTTTGACTATTTGTGTATTCACGCGCTTACAGCCAGCGGCTGATATG 1404
Db 1416 GTGCTTCAGAGTTTAATCTGTTGCTCATCCGAGCTATGTGTACAGGCGCTGCAATC 1475
QY 1405 GCGTTCCGCGGTAGCTCAAGTGTGTGTGTAACCACTGTCAACCCGCACTGGCTAT 1464
Db 1476 CCACTCAGATCAACCATTTAAGATCTTTTGGCAAGAAAGCAAGC---GAGAGTTAATA 1532
QY 1465 ACAATGCCGATGTGAGCATCTGCTGTGCTTACAGCAAGAGTGAAGATGACTTCA 1524
Db 1533 CTTGTTCTGACGCGGTGCTGACGAGGATGTGTACAAAGAGAGAGATGTGGAG 1592
QY 1525 GCGGAGAGATCAGAAAGCCCTGACCTTGTCTCTGTAATGAAAAAGAGACGCTA 1584
Db 1593 GCCAGCTGACTAAGAAAGCTTACCAACCTCGTCCCATCGCGCGGCGAGACGGGA 1652
QY 1585 GAGGCAAGATGACCTCATATATGCAAG---GAGAGAGAGATGAGGTGAGACA 1638
Db 1653 GTCTGCACTGCACTTTGAAGAACTAGAGATCAACGAGGTGCAAGGGGTGTGAAACA 1712
QY 1639 CCAGAGGACACATCAGGGGTGACAAAGTTACCGAGCGAAGAGATGGGTCTTACGCT 1698
Db 1713 CTTGCGACGCGGTGAAAGTCAACGCGACAGCGAAGCATGTAATGAAATTTAGCTA 1772
QY 1699 ATACTTTCAACCCAGCGGTATGTAATGTAAGAAATGCGGTGTATCCACCAATGGCG 1758
Db 1773 GTTCTGTCGCCGAGACCGTGTCTCAAGAGCTCAAGTTGGCCCCGCTACCTCTAGCA 1832
QY 1759 GAACAAGTCTGTATGATCTCAAAAGTGGGCAAGGAATACAAAGTGAAGCATAC 1818
Db 1833 GAGAGGTGAAATTAACATTAACGTAGGGCGGCGGTACCAAGTGAAGGATAT 1892
QY 1819 CAGGTAAGTCTATTTGATCAGAAAGGAGCGCGGCTCTGTTCAAGATTCCAGGCAATG 1878
Db 1893 GACGCGAGGTCTCTAATCACTGTGTGATCGGCAATTCGGGTCTCTGAGTTTCAAGCTT 1952
QY 1879 AGTGAAGCGGTACGATGTTTCAACGAGAGGAGGATTCGTAACAGATACCTGACAC 1938
Db 1953 AGCGAGAGCGCACTATGTGTGAACGAAAGGGGTTGTCAACAGGAACTATACAT 2012
QY 1939 ATCCGATCAACGAGAGCGCTAAACCTGACGAAGATCTATTAAGCTGTAAAGCT 1998
Db 2013 ATTGCGTTTCAAGGACCGTGTCTGTAACACGAGAGGAGAACTACGAAGATGAGACT 2072

QY 1999 CAGGACAGACTCAGAAATACGTTCTTGATATTGACGACGAAAGTGTGTTAAGCAGAA 2058
Db 2073 GAAAGAACTAGCCCGGAGTACGTTGTCGAGTGTAGTAAAGAAAGTCTGCTCAAGAGAG 2132
QY 2059 GACGAGGTCCCTTGTGCTTAACCGGTGATCTGTGATATCAACATTTACAGATTGGC 2118
Db 2133 GAAGCGTGGGTTGGTGTGTGGAGAGCTAACCAACCCCGTTCATGATTTGGC 2192
QY 2119 TAGGAGTCTCAAGACAGACAGCAGCACTCAACAAAGTCCCAACCATCGAGTCTAT 2178
Db 2193 TACGAGGCTGAAGATCAGGCGGTGCGACCATATTAAGATACAGTATGAGAGTCTTT 2252
QY 2179 GAGTGCAGGTTCAAGTAAATCTGAAATCATCAAAAGCCGTGTGACTAAGAAAGATCTG 2238
Db 2253 GGGGTTCCGGGATCAGCAAGTGTGTATTAAGAGCTCGTGACCAACAGATCTG 2312
QY 2239 GTTGTAGTGCAGAAAGAAAGAAATCTGCGCAAAATCATCAGGATGTAAAGAGATAGA 2298
Db 2313 GTCAACGCGGCAAGAGAAAGAACTGCGAGAAATATTTAAACGACGTGAAGAAAGCACCGC 2372
QY 2299 CGTATGATGTGCTGTAGAGCTGTGATTCAGTGTCTTAAATGGGGTTAAGCACCCC 2358
Db 2373 GGGAGGGGACAGGTAGGAAACAGTACTCATCTGCTAAGCGGTGTCTGTGCGC 2432
QY 2359 GTTAACACTCTGTACATTGATGAGGCAATTTGCTGCCATGACAGGAGCGTGTGGCACTG 2418
Db 2433 GTGACATGTCTATATGTGAGAGAGCTTGTGCTTGCCATTCGGGTACTGTGCTGGGCTA 2492
QY 2419 ATTGCAATGCTCAAACT---AGAAAGTGTATTTGTGCGGAGACCCAAACAATGCGC 2475
Db 2493 ATTGCTCTTGTAAACCTCGAGCAAGTGTGTATGACGAGACCCCAAGCAATGCGGA 2552
QY 2476 TTCTTTAAGATGATGTGCTGAAGATCATTTTAACCATGACATATGACATGAAAGTATC 2535
Db 2553 TTCTTCAATATGATGACGCTTTAGGTGATTTCAACCAACATCTGACTGAGATGTGT 2612
QY 2536 CATTAAGACATCTCTAGAGGTGACACAGACTGTAAACCGCATGTCTCCACGCTTTC 2595
Db 2613 CATTAAGATATATCACAACGTGACAGCGTCAAGTCAACGCGCATCGTGTACGTTGCAC 2672
QY 2596 TAGCAACAGGATGAAGACGTTTAAACCATGTGCTATTAATCATATGATACACA 2655
Db 2673 TACGAGGCAAGATGCCACAGCCACCGTGCACAAACCCATTAATCAATAGACACACA 2732
QY 2656 GGGACCAAAAGCGGCAAAAGATGATCTGATTTAATCTGTTTCAAGATGGGTGAAA 2715
Db 2733 GAGACAGCAAGGCCCAAGCAAGAGATGTGTATTAATGATCTTCCAGGCTGGCAAG 2792
QY 2716 CAGCTACAGATTGACTCAAAATCAAGAAATCATGACTGCGGTGATGCAAGACTT 2775
Db 2793 CAGCTGCAATTGACTACCGTGTGACAGAGATCATGACAGAGCATCTCAGGGCTC 2852
QY 2776 ACGCGAAAGCGTTTATGCTGTGAGTACAAAGTCAACGAGATCTACTGCGAG 2835
Db 2853 ACCCGCAAGGGGATATCGCGTAAAGCAGAAAGTGTATTAATATCCCTGTATGCCCC 2912
QY 2836 ACTTGAGACGTAAGCGTGTACTTACACGACAGAAAGCAATGTTCTGGAAGCG 2895
Db 2913 GCGTGGAGGACGTAATGTATCTGTGACGCGCACTAGATGAGTGTGTGAAAGCG 2972
QY 2896 CTAGCTGTATCCCTGATTAAGACACTTACAGTCAATATCCCGGGGATTTACGCGCT 2955
Db 2973 CTGGCGGCGATCCCTGATTAAGGTCTTATCAAAAGTTCCACAGGGTAACTTTACGGCC 3032
QY 2956 TCATTGACAGCTGTGACGCGGAAACAGACGCAATTAAGCAAGCGCTTTTGATTAAGCG 3015
Db 3033 ACATTGGAAGATGTGCAAGAAAGAACAGCAAAATTAAGAGTGTATTAAGAGCGCGCT 3092
QY 3016 CAGACAGCTATGTGTTCCAGAAATTAAGTGAAGTGTGCGGGGAGAGGCTTTAGAGCA 3075
Db 3093 GCGCTGTGACGCGTTTCCAGAACAAAGCAAGCTGTGTGTGTGAAAGAGCTGTGTCT 3152
QY 3076 GTCTTGGCCACGCGCAACATTTGTGTGTGACGAGACAGCAGTGGGAGAC---GTTCACCCA 3132

Db 3153 GTCCGAGCACTGCGGAAATGATGACAGCAGAGAGTGAGCACCATAATACAGCA 3212
Qy 3133 TTCAGCAATGACAGAGCGTAATCACTGAAATGACAGCACTTTCTTTGACCAAGTTC 3192
Db 3213 TTTAAGAGAGACAGAGCTTACTCTCCAGTGGCTTGAATGAATTTGACCAAGTAC 3272
Qy 3193 TTTGAGTAGACCTGGAAGAGTGGTATTTTTCGCTCTCAACCGTGCACCTTACTTACAG 3252
Db 3273 TATGAGATTGACCTGGAAGAGTGGTATTTTTCGCTCTCAACCGTGCACCTTACTTACAG 3332
Qy 3253 GATCAGCACTGGGATTAATCTGCGCAGAGGAAAGCATGTATGGCTTAATAGAGAGTGA 3312
Db 3333 AACACCACTGGGATTAATCAAGACTGCTGGAAGATGTATGATTAATCCGGAAACAGCT 3392
Qy 3313 AAGAGTTGTACCGGCAATATCCGTCATCAAAAAGCGGTGACACAGCGAGGTAGT 3372
Db 3393 GCCAGCTGGAAAGCTAGACATACCTTCTGAAGGGGAGTGGCATACGGGCAAGAGCA 3452
Qy 3373 GATATAGGAATATATACATCAAGAGCTACTCTCAACAAATTAATGGTTCATTAAT 3432
Db 3453 GTTATCGCAGAAAGAAAATCCACCGCTTCTGTGCTGAGACAAATGTAATCTTATCAAC 3512
Qy 3433 CGCCGTTGCCCTCTGTTGATCGTTGACCAAAAGACAGGGTACAACTGATCAAGC 3492
Db 3513 CGCAGGCTGCCGACGCGCTGGCTGAGTACAAAGACGTTAAAGGAGTGGGTGAG 3572
Qy 3493 GGATTCCTATCTAATGAAGGCAATCTGTGTGTGATCGGGATCTTATCAGCAT 3552
Db 3573 TGGCTGTCAATTAAGTAAGAGCTTACCACTGTCTGTGTGATGATGATCAACCTGGCT 3632
Qy 3553 CCAGGAGAAAGTATGAGTCCATGGGTCCATTGGCCCATTAATAC-----CATCAGG 3603
Db 3633 TTGCTCCACGAGAGGTCACTTGTGTGTACCTGATGTACACAGCGCCGATAGTGC 3692
Qy 3604 TGTGATCTGATTTGGGAATACCTAGCAATGTGGTAAATATGATTAATCTTGTCAAT 3663
Db 3693 TACGACTTAAGTTTGAAGTACTGCGGCTAGCGCGGAGGTGATGACTTGTGTGAAAC 3752
Qy 3664 GTTAGAGCCCGTACAGGAACCATCACTACCAAGTGGAGATACGCTATCCACAC 3723
Db 3753 ATTCAACAGGAATTCAGATTCACCACTACAGAGTGTGTGACACGCGCCATGAAGTGC 3812
Qy 3724 AGCATGCTAAGCTGTAAGGCTGTCCACCACTGAACATGTGGGGAACATGTGTGCTATA 3783
Db 3813 CAGATGCTTTGGGAGATGCGTACGATCTTAACCCGCGCATCT--TGAAGA 3869
Qy 3784 GGGTATGGGCTTGTGATCGCGCAACGAGAAATATCATCATGCGGTGCAAGCTCATTT 3843
Db 3870 GCTTACGGAATACCGGATTAATCAAGCAAGCGGTGTTTCTCTTAAGCAAGAAAGTTC 3929
Qy 3844 AGGTTTACCGGTGTCTGTACAGCTTAAGAACTGCCGAAATATGAGAGTCTTCTTGGT 3903
Db 3930 TCGTGTGAAGAGTGTGGCGCGGATGTGTACACAGCATATACGAAGTGTCTTGTGCTG 3989
Qy 3904 TTCTTGGCAAGGAGCAAGGCAACCAACATATGACAGAGAGTGGTGTAGTGGT 3963
Db 3990 TTCTTCAACTTTGACAAAGGAAAGAGACCTCTTACGCTTACACAGATGATACCAAGCTG 4049
Qy 3964 GACAACTTATCAAGGCTCAACAGGTACAGGAGGAGGAGCTCAAGCTACAGAGTG 4023
Db 4050 AGTGCCTGTATGCGGAGAAAGGATGACACAGCGCGGTGTGACATCTCAGAGAGTT 4109
Qy 4024 ATCAGAGGTGACATTAAGCAAGCGCTGACCAAGCTATGTTAATGTGCTAATAGCAAA 4083
Db 4110 AAGAGAGCAGATATAGCCAGTGCACAGAGCGCTGTGTGTATAGAGAGTACCCCGT 4169
Qy 4084 GGTCAACCAAGTTCCGAGGTGTGCGGTGACCTGTACCCGAAATATGCGCGGCTTTTAT 4143
Db 4170 GGAATCTAGGAGATGCGTATGACAGGGCTGTGGAGAAATATGCGCTTAAAG 4229
Qy 4144 AGACAGCAATATGCTGTGCGGACGCGTGAAGCTTGTGAAGACGAACCG---CTCATCATTA 4200

Db 4230 GGAGCAGCAACACAGTGGGCAATTAATAACATCATGTGCGCTCGTACCCGTCATC 4289
Qy 4201 CATGCTAGAGACCCAAATTTTCTTAAGATGCCGGAACCGGAGGGGACCTTAAGCTCGA 4260
Db 4290 CACGCTAGAGGCTTAATTTTCTGTCCACGACTGAAGGGAAGGGGACCGCAATTTGGCC 4349
Qy 4261 GCTGCTCATAGAGCATATGCTCCATCGTCAACGCTGAGCGGATTAACAAAATATCAGTA 4320
Db 4350 GCTGTCTACCGGGCAGTGGCCGCGAATTAACAGACTGTCACTAGACAGCTTACCATTC 4409
Qy 4331 CCGCTACTGTCAACCGGCATCTTATTTGTGTGCAAAATTCAGATGATCATATTCAT 4380
Db 4410 CCGCTGTGTCCACAGAGTGTTCAGCCGCGGAAGATAGGCTCAGCAATCTCCCTCAAC 4469
Qy 4381 CACCTGTCACTGCTTTGACATACAGGATCCGATGTCAACATATATTTGCTGGATTA 4440
Db 4470 CATCTATTCAAGCATATGAGACGCAAGACCTGACGTGACCTTACTGACAGACAA 4529
Qy 4441 CAATGGAGACCAAGATATCGAGGCAATTCACCGCAAAAGAAAGCGTCAAAATTCGAT 4500
Db 4530 AGTTGGAGAAAGAAATTCAGGAAGCCATTCATGACAGAGAGCGCTGTGAGTGTCAAT 4589
Qy 4501 GATGCAACCAAGTATGATGATCTGTCAGAGGTCCACCCAAACAGCTCTTTGGCAGGC 4560
Db 4590 GATGACGTGAGCTGACACAGACTTGGTGAAGTGCACCGGACAGCAGCCTGGTGGT 4649
Qy 4561 AGACAGGTTACTCCGTAAATGAGGCAAGTTGATATCTGCTGGAAGTATCAGATTC 4620
Db 4650 CGTAAAGGCTACAGTACACAGCGGTGCTGTACTGTACTTGAAGTATCGAAATTC 4709
Qy 4621 CATCAACCGCCAAAGACATTCGCCGAATTCATGCAATGTGGCCCAACAAATCTGAGCT 4680
Db 4710 AACCAAGCTGCTATTTGATATGAGCAGATATCTGAGCTTGTGGCCAGACTCAGAGGCA 4769
Qy 4661 AATGACAGATTTGCTGTATCATCTGTGGGAGAGATGTCCAGATCCGCTCCAAATGC 4740
Db 4770 AACCAAGATATGCTATATGCGCGTGGCGCAACAAATGACATCAGATCCAAATGT 4829
Qy 4741 CCAGTATGAGATCAGAGCGCTGTCTCCTCACTACACACTTCCATGCTGTGTAATTC 4800
Db 4830 CCGGTAACGATTCGATTCATCAACACTTCCAGAGAGTGCCTGTGCTGCGCTAC 4889
Qy 4801 GCTATGACGCTGAGCGGCTATACAGTGTGGCTCTGCGAAGAAAGACATTCGCCGTA 4860
Db 4890 GCATGACAGAGAGAGATGCGCGGCTTATGATGATCACCAGTTAAAGCATGTGCT 4949
Qy 4861 TGTCTATCATTTCTGTGCGCAAGTATCAGAGCGTGCAGAACTACAGTGCAGC 4920
Db 4950 TGTCTATCTTTTCCCTCCGAAATATCAGTATGATGGGGTGCAGAAAGTAAAGTGCAG 5009
Qy 4921 AAACCAATCTGTTTTCAGGCGTGTATCACCGGCTGTATCACCCAGAGATAGCGGGA 4980
Db 5010 AAGGTTCTCTGTTTCAGCCGACGATACCTTCAAGTGTATGTGCGGAGATATCCCA 5069

RESULT 15
AAK78129
ID AAK78129 standard; DNA; 8010 BP.
XX
XX AAK78129;
XX
XX 20-AUG-1999 (first entry)
XX
XX
XX Semliki Forest virus replicon fragment.
XX
XX Alphavirus; replicon; expression vector; heterologous splice site;
XX RNA splicing; viral RNA replication; antigen; vaccination; therapy; ss.
OS Semliki forest virus.
XX
XX W09925859-A1.
XX
XX 27-MAY-1999.
PD

XX 13-NOV-1998; 98WO-CA01065.
XX 14-NOV-1997; 97US-0065793.
XX (CONN-) CONNAUGHT LAB LTD.
XX Parrington M;
XX WPI; 1999-385140/32.
XX
XX New vector (EV) used to express an antigen for vaccination, e.g.
XX against human immune deficiency virus
XX
XX Disclosure; Fig 3A-F; 60pp; English.
XX
XX This invention describes the construction of a novel expression vector
XX (EV) containing an alphavirus replicon having a heterologous splice site.
XX The expression vector contains (i) a DNA molecule (i) complementary to
XX at least part of an alphavirus RNA genome, and (ii) inserted into a
XX region of (i) that is not essential for its replication, a heterologous
XX DNA (ii), under control of a promoter. (i) includes at least one
XX heterologous splice site (HSS) to prevent aberrant RNA splicing of the
XX alphavirus and is the complement of the complete alphavirus genomic
XX region essential for replication of viral RNA. The expression vector is
XX used to express (ii) in humans or animals, e.g. to express an antigen
XX for vaccination (against human immune deficiency virus) or to produce a
XX therapeutically active protein or peptide. Introducing an HSS: (1) makes
XX it more likely that any splicing will occur at this site, rather than at
XX a cryptic splice site; (2) restores function of the alphavirus when
XX eliminated, and (3) may improve transport of RNA from the nucleus.
XX
SQ Sequence 8010 BP; 2213 A; 2036 C; 2158 G; 1603 T; 0 other;

Query Match 29.5%; Score 1473.8; DB 20; Length 8010;
Best Local Similarity 57.9%; Pred. No. 0;
Matches 2779; Conservative 0; Mismatches 1982; Indels 36; Gaps 8;

Qy 214 ATTGAAGTGGCCGCTGACATGACATTCACATCCGCTATCATTCGTCCT 273
Db 1 ATCGCAGTGGCCCTTCAGAGAAATGATGTCTACGCAAAATCCACTGCTATGCTCT 60
Qy 274 ATGTAAGCCGCTGAAGACCCGAGACACTACACGGTATGAGAAAGCTTAAGAAAA-- 331
Db 61 ATGGCAGCGCAGAAAGACCCGAAAGGCTGATAGCTACGAAAGAAATCGCAGCGGCC 120
Qy 332 ----GTGACATTACCGACAGAAACATAGCTCTTAAGCGCGAGACTGTGGAAGTCATG 387
Db 121 TCCGGGAAGGTGCTGATAGAGATCGCAGAAAAATCACCGACTGACACCGTCATG 180
Qy 388 TCAACACGACGACGAGAGACTCCATCTCTGTATGACACAGACGGCAGCTGATAGTAC 447
Db 181 GCTACGCCAGCGCTGAATCTCTTACCTTTTCTGCTGACATACAGACGTACAGTGTGACG 240
Qy 448 TTGGAAGTGAAGATATACCAAGATGTGACAGCTGATGACCGACATCAATCTAC 507
Db 241 GCACCCGAAAGTGGCGATATACAGGAGCGTATGCTATGACCAACATCGGTGAC 300
Qy 508 CACCAAGCGCTTAAAGAGATTAGCAATTTACTGATAGGCTTTGACAGACCCCTTTT 567
Db 301 CATAGGCGATGAAGAGGTGTGAGAAACGCGCTATTGATGGGTTTGAACCAACCCCGTTT 360
Qy 568 ATGTACAAAACATGAGCGATTTCTACCTTACTTACACAGAACTGGGCTGAGAGAGA 627
Db 361 ATGTTGACGGCTAGCAGCGCGGTATCAACTACGCAAACTGGCGCAGAGAGAG 420
Qy 628 GTATTGAGACGCTAACATTTGGCTCGTAACTCAGATCTTCAAGAGAGAGCGCTTGA 687
Db 421 GTGTTACAGCGCAGAAATAGACTGTGTGACAGATCTTGAATGAGAGAAAGATCGGC 480
Qy 688 AAATCTCAATCTTAGAGAGAGAGGCTCAACCTACTTAATAAGATCATTTCTCGTT 747
Db 481 AAATGTCATTTCTCCGAGAGAAACATTTGAACCTTGACACAGATCATTTCTCGGTA 540

Qy 748 GGTTCACAAATCTACACAGAGATAGATCACTGTATAGTGGCTGCACTTCCAAAGCTG 807
Db 541 GGATCTACATTTGACATGAGAGAGAAAGCTACTGAGAGCTGGCACTTACCTCCGTA 600
Qy 808 TTCCACTTGAAGAGAAAGCTAACTTACAGATGATGAGGAGCAATGTCTAGCTGTGA 867
Db 601 TTCCACTTGAAGAGAAAGCTTAACTTACCTGATGAGGAGATACATGATCATGTGAA 660
Qy 868 GGGTACGTCATCAAAAGATTAACATCAGCCAGAGACTATACGTTAAGTTGAGAACTTG 927
Db 661 GGGTACGTCATTAAGAAATCACTATGTGCCCCGCTGTACGTTAAACGTTAGGGTAC 720
Qy 928 GCGTCCAAATGATCGCAGGGGTTTCTTGTAGTTGCAAAAGTACAGATACGCTGCGCGC 987
Db 721 GCCGTACGATACACGGGAGGAGATTCCTAGTGTGCAAGACACAGACATCTGTAAGGA 780
Qy 988 GAGAGGTTTCTTTTCTGTGTGTGATGATATGACCAACCACTTGGCATCAGATACA 1047
Db 781 GAAAGAGTCTATTCCTGTATGACCTTACGTCCTCAACATCTGTGATCAATGACT 840
Qy 1048 GGGATTTGGCACTGACGTTAGTGTGATGACGCAACAAAACATTTGGTGGCTAC 1107
Db 841 GGCATCTACGACGACGCTACACCGAGAGACGACAGAACTTTAGTGGGATTAAT 900
Qy 1108 CAAGATGTCGTCATGTTAGAGACGCAAAAGAAATCTAACAACATGACAGAACTATCTA 1167
Db 901 CAGAGATGATTTGTAACGAGAAACACAGCAAGAACTAACAAGATGAAAGTATCTG 960
Qy 1168 TTACCACTGTGCGCCAGGGGTTTTCAGGTGGCGGTGAACATCTGCCGACTTGAC 1227
Db 961 CTTCGATGTGGCGCTGCGCATTTAGCAAGTGGCGAGGAATCAAGCAGACCTTGAT 1020
Qy 1228 GAGCAGAAAGATCTAGGGGTGGGAGCGCACTTTATCTATGGCTGTGCTGGCTTTC 1287
Db 1021 GATGAAAAACCTCTGGGTGTGTCAGAGAGTCACTTACTGTGCTGTGTTGGCAATTT 1080
Qy 1288 AAGACCAGAAATCAATCATCTACATCAAGAGGCTGTACGCAAAATTAAGAAAGTA 1347
Db 1081 AAAACAGAGAAATGACACACATCTACAGAAACACACACAGCAAAATAGTGAAGTG 1140
Qy 1348 CTTGCGCTCTTGAATCTATTTGTGATTCACGCTTACAGCCAGCGGCTCGATATGGC 1407
Db 1141 CTTTCAAGATTTAATCTGTCTATCTCCAGCTTATGCTACAGGCTCGCAATCCA 1200
Qy 1408 TTCCGCGTATGCTCAAGCTGTCTCTTGAACCAACTGTCAAAACCCGACCGGCTATTA 1467
Db 1201 GTCAGATCAGCATTAAGATGCTTTTGGCCAGAAAGACCAAGCGAG--GTTAATACT 1257
Qy 1468 ATGACCATGTGAGCATCTGCGTGTCTACAGCAAGAACTGAAAGTGGCTGACGCG 1527
Db 1258 GTTCTCAGACCGTGTGACGACGAGGATGCTGAACAAAGAGAGAGAGTGGAGGCC 1317
Qy 1528 GAAGATACAGAGAAAGCTTGCACCTCTGCTCTGTAATTAAGAAAAAGACCGTACG 1587
Db 1318 GAGCTGACTAGAGAAAGCTTACCAACCTCTGCTCTGCTGCGCGCGAGAGCGGAGTC 1377
Qy 1588 GCAGAAATTAACCTCATTTAGCAAGAG-----GCAGACAGGTAGCGTGAAGACCA 1641
Db 1378 GTGACGTGACAGTTGAAGAACTTAGATATCAAGCAGAGTACAGAGGTGCGGAAACCT 1437
Qy 1642 GCAGGACATATCAGGGGTGCAAGTTACCGACGAGAGAGAAAGATTGGGTCTTACGCTATA 1701
Db 1438 GCGAGCGCGTTGAAGATCACGCGACGACCGCAAGCAGCTACTTAGAAATTTAGTACT 1497
Qy 1702 CTTTACCCGAGCGGATTAATTAATGTAAGAACTGGCGTGTATCCACCCATTTGGCGGAA 1761
Db 1498 CTGTCCCGAGACCGTGTCTCAAGAGCTCAAGTTGGCCCCCGGACACCTCTAGCAGAG 1557
Qy 1762 CAAGTACTGTGATGATCAAAAGTGGGCGAGGAGATACAAAGTGTGAGCCATACAC 1821
Db 1558 CAGGTGAAAAATTAATTAACATTAACGAGGCGGCGGTTTACCAAGTGTGACGATATGAC 1617

QY 1822 GGTAAAGTCATTGTAACGAAAGGACGCGGCTCCCTGTTCAAGACTTCCAGGCATTGAGT 1881
Db 1618 GCGAGGGTCTTACTACTAGTGGATCGGCGATTCGGGTCTCTGAGTTTCAAGCTTTGAGC 1677
QY 1882 GAGACGCTACGATCGTTTTCACGAGAGGAGTTCTGTAACAGATACCTGCACCAATC 1941
Db 1678 GAGACGCGCACTATGCTGTACACGAAAGGAGTTCTGCACAGGAAATATATACATATT 1737
QY 1942 GCAATCAACGAGAGCGCTTAACACTGACGAAGATCTATTAAGCTGTAAAGACTCG 2001
Db 1738 GCCGTTCACGACCGCTCGCTGAACACGAGAGAACTACGAGAAATCAGAGCTGAA 1797
QY 2002 GACACAGACTCAAAATACGTTCTTCATTTGACGACGAAGTGTATTAGCGAGAAAGC 2061
Db 1798 AGAACTGACGCGCGATACGTTGTTGACGATTAATAAAATGCTGCTGAAAGAGAGGAA 1857
QY 2062 GCAGGTCCCTTGTGCTTAACCGGTATCTGATAGATCCACCATTTTCAAGATTGCGTAC 2121
Db 1858 GCGTCGGGTTGTGTGTGGAGAGCTAACCAACCCCGTTCATGMAATTCGCTTAC 1917
QY 2122 GAGAGTCTCAAGACACGACACGACAGCCTCAAAAGTCCCAACCATCGGAGCTTATGGA 2181
Db 1918 GAAAGGCTGMAAGATACAGCCCGTCGCGACCATTAAGACTACAGTGTAGAGTCTTTGG 1977
QY 2182 GTCCAGGTTCAAGTAAATCTGGAATCATCAAAAGCGCTGTGACTAAGAAAGATCTGGT 2241
Db 1978 GTTCCGGGATGAGCAAGTCTGCTATTATTAAGAGCTGTGTGACCAACACAGATCTGGT 2037
QY 2242 GTGAGTGGAGAGAGAAAGAACTGCGCAGAAATCATCAGGAGTGTAAAGAGATGAGAGCT 2301
Db 2038 ACCACGCGCAAGAGAGAACTGCCAGGAAATGTTAAAGAGTAAAGAGACCGCGGG 2097
QY 2302 ATGAGTGTGCTGCTAGAGCTGTGATTCAGTCTTCAATAGGGGTTAAAGACCCCGT 2361
Db 2098 AAGGGGCAAGTAGGAAACAGTACATCTGCTTAACGAGGTGTGTGTCGG 2157
QY 2362 AACACTGTATCATTTAGTAGAGCATTTGCTGCTGCAATGACGAGCGCTGTGCACTGAT 2421
Db 2158 GACATCTTATATGTGAGCAGAGCTTTCGCTTCCATTCGGTATCTGTGCTGCTTAAAT 2217
QY 2422 GCCATGCTCAAACT--AAGAAAGTGTATTGTGCGGGAGCCCAAAACATGCGGCTTC 2478
Db 2218 GCTCTTGTAAACCTCGAGCAAGTGTGTATCGGAGACCCCAAGCAATGCGGATTC 2277
QY 2479 TTTAATCATGATGCTGCTGAAAGTATACATTTTACATGATATGAGTGAATGAT 2538
Db 2278 TTTCAATATGATGAGCTTAAGGTGAATTTCAACCAACATCTGATGAGTATGAT 2337
QY 2539 AAAAGCATCTCTAGAGGTGACACAGACTGTAAACCGCATGCTCCACGCTCTTCTAC 2598
Db 2338 AAAAGTATATCCAGAGCTTTCACGCGTCAAGTCAAGGCTATGCTTACGTTGACATC 2397
QY 2599 GACAAAGCAATGAAAGCGTTTAAACCATGTGCTGATTAATATCATATGATACACAGGG 2658
Db 2398 GAGAGCAAGATGCGACGACCAACCCGTCACAAACCATATATAGACACACAGGA 2457
QY 2659 ACCAACAAGCCGCAACAAGATGATCTGATTTCACTGTTTACAGAGATGGGTGAACAG 2718
Db 2458 CAGACCAAGCCCAACCCAGAGACATCGTGTAAATGCTTCCGAGGCTGGCAAAAGAG 2517
QY 2719 CTACAGATTGACTACAAAAATACAGAAATCATGATCGGCTGATTCGCAAGAGCTTACG 2778
Db 2518 CTGAGTTGAGATACCGTGGACACGAAAGTCAATGACAGACAGCATCTCAGGAGCTTAC 2577
QY 2779 CGGAAAGCGTTTATGCTGTCAAGTAAAGTCAACGAAATCACTTACTGCGAGACT 2838
Db 2578 CGCAAAAGGGTATAGCGCGTAAAGGACAGAGTGAATAATCCCTTGATCCCTCG 2637
QY 2839 TCTGAGCAGTGAAGTGTATTACTTACAGCAAGAAACGATGTCTGGAAGAGCGCTA 2898
Db 2638 TCGGAGCAGTGAATGATCTGACGCGACTGAGAGATAGGCTGTGGAAGAAAGCGCTG 2697
QY 2899 GCTGTGATCCCTGATAAAGACACTTACAGCTAAATATCCGCGGATTTTCAAGCTTCA 2958

Db 2698 GCCGCGATCCCTGATTAAGGTCTTATCAAACTTCCACAGGGTAACTTAAAGCGCCACA 2757
QY 2959 TTGACGACTGCGACCGGAAACAGACCGCATTAATGACCGCTTCTTGAATGAGCCGAG 3018
Db 2758 TTGAAAGATGGCAAGAAACAGCAAAATTAAGAGTGTATGAAGACCGGCTGCG 2817
QY 3019 ACAGCTGATGTGTTCCAGAAATAGGTGAACGTCTGCTGGCGAAAGCTTTAGAGCAGTC 3078
Db 2818 CCGTGGACGGTTCACAAACAAAGCGAAGCTGTGTTGGGAAAGAGCTGTGCTGTC 2877
QY 3079 TTGGCAGCGCCAACTTGTGCTGACGAGACAGAGTGGAGAC--GTTCACCCATTC 3135
Db 2878 CTGACACTGCCGGAATCAGATTGACACAGAGAGGTGAGACACATTAATTAACACATTT 2937
QY 3136 AAGCATGACAGAGCTATCACTGAAATGACACTGAACTTCTTTTGACACAGGTTCTTT 3195
Db 2938 AAGAGGACAGAGCTTACTTCCAGTGTGGCTTGAATGAATTTTGACCAAGTACTAT 2997
QY 3196 GAGTGAACCTGACAGTGGGTTATTTCCGCTCCTACCGTGCACCTTACTTACAGGAT 3255
Db 2998 GAGTGTGACTGACAGTGGCTGTGTTTCTGCGCGAAGGTGTCTCTATTTACAGAAC 3057
QY 3256 CAGACTGGATTAACCTGCCAGGAAAGACATGATGGGCTTAATGAGAGGTACAAAG 3315
Db 3058 AACCACTGGATTAACAGACTGTGGAGAGATGATGATTTCAATGCCCAACAGCTGCC 3117
QY 3316 GAGTTGTACGCGCATATCCGTGATCACAAAAGCGTTTGACACAGGAGTACTGAT 3375
Db 3118 AGGCTGAAGGTAGACATACCTTCTGAAAGGGGAGTGGCATCCGGGCAAGCGACAT 3177
QY 3376 ATTAAGAAATTAATCATCAAGAGATACCTCCCAATTAATGCTTCCATTAATGCG 3435
Db 3178 ATCCGAAAGAAATTAATCAACCGTTTCTGTGCTGACAAATGATTTCTATCAACGC 3237
QY 3436 CGGTTCCCACTCGTTGATGCTTGACCAACAAGACAGGTACAACTGATCACAGCGGA 3495
Db 3238 AGGCTGCCGACGCGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 3297
QY 3496 TTCTATTAATGAAGAGGCAATCTGTGTTGATGATGATGATGATGATGATGATGAT 3555
Db 3298 CTGCTCAATTAAGTAAGAGGTATCCACGCTCTGCTGTGATGATGATGATGATGATGAT 3357
QY 3556 GGGAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3606
Db 3358 CCGACGACGAGGTACCTGTTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 3417
QY 3607 GATCTGATTTGGAAATACCTAGCCATGTCGTAATTAATGACATTAATTTGTCAATGT 3666
Db 3418 GACCTAAGTTTAAAGATGCTGCGGCTGACGCGGAGGTTCGACTTGTGTTGTAACAT 3477
QY 3667 AGGACCCCGTACAGAACATCACTACCAAGTGTGAGATCAAGCTATCCACACAGC 3726
Db 3478 CACACGGAATTAACAATCAACACTACAGAGTGTGTCACACGCGCATGAAGCTGAG 3537
QY 3727 ATGCTAAGTGTAAAGCTGTCCACCACTGAACCTGCGGAAACATGTTGTGCTATGAGG 3786
Db 3538 ATGCTGGGAGATGCGTACGACTGTAACCCGCGGCACTCT--TATGAGACT 3594
QY 3787 TATGCTTGTGCTGATCGCGCAACCGAATTAATCACTGCGGTGACAGCTCATTTAGG 3846
Db 3595 TACGATATGCGCATTAATTAATGAGGAAGCTTGTCTCTTAAAGCAAAAGTTCTCG 3654
QY 3847 TTTACCGGTGTCTGACGCTTAAGAACATCTCCGAAATTAAGAGTTCTTCTGTGTC 3906
Db 3655 TCTGCAAGAGTGTGCGCCGATGTTGTCTACACAGCATTAAGAAAGTCTTGTGCTGTC 3714
QY 3907 TTGCGCAAGGACCAACGCAACACATACACAGGACACACTCGGTGTGCTGATAC 3966
Db 3715 TCCAACTTTGACAGGAAAGAGACCTCTACGCTTACACAGATGAATACAGAGTGAAT 3774
QY 3967 AACACTATCAAGGATCAACAGGTACAGAGGACAGGAGAGCTTCCAGGCTTACAGAGTATC 4026

Db 3775 GCCGTGATATGCCGGAAGACCATGACACGCGCGGTGTGCAACCATCTTACAGATTAAAG 3834
QY 4027 AGAGGTACATTACCAAGAGGCTGACCAAGCTATCGTTATGCTGCTTAATAGCAAAAGT 4086
Db 3835 AGAGCAGACATAGCCACGATGACAGAAAGCGGCTGTGTAAAGCAGCTAACGCCGTGGA 3894
QY 4087 CAACCAAGTTCCGGAGTGTGCGGTGACCTGTACGAAAATGGCCGGCTGCTTTTGATAGA 4146
Db 3895 ACTGTAGGGGATGTGCGTATGTAGAGGCCCGTGGCGAAAGAAATGCGCTTAAGGA 3954
QY 4147 CAGCCAAATAGCTGTGCGGACGGCTAGACTGTGAAGCAGAACCG--CTCATATACAT 4203
Db 3955 GCAGCAACACACGATGGGACATTTAAACAGTCAATGTGCGGCTGTACCCCGTATCCAC 4014
QY 4204 GCTTAGAGACCCATTTTCTTAAGATGCCGAACCGAGGGCGACTTAAGCTGCCAGCT 4263
Db 4015 GCTGTAGCGCTTAATTTCTTGCCACGCTGAAGCGGAAGGGGACCGGAATTTGGCGCT 4074
QY 4264 GCCTACATGAGCTAGCGTCCATGCTCAACGCTGAGCGGATTACAAAATATCAGTACCG 4323
Db 4075 GTCTACGGGGAGTGGCCGCCGAAGTAACAGACTGTACCTGAGCAGGTAGCCATCCG 4134
QY 4324 CTACTGTCAACCGGCTATCTATTGTGTGCGAAAGATGATGCAATTCATTCATCAC 4383
Db 4135 CTGCTGTCAAGAGAGTGTTCAGCGCGGAGAGATAGGCTGCAGCAATCCCTCAACCAT 4194
QY 4384 CTGTTCACTGCTTTCGACATACGAGATGCCGATGTCACCATATATGCTTGATTAACAA 4443
Db 4195 CTATTCAAGAAATGAGCGCACGACGATGACGATCTACTGACAGACAAAGT 4254
QY 4444 TGGGAGACCAAGATATGAGGCGATTCAACGCAAGAAAGCGTGAATTCGAGATGAT 4503
Db 4255 TGGGAGAAAGAAATTCAGGAAGCATTAAGATGAGAGCGGTGTGAGTTGCTCATATGAT 4314
QY 4504 GACAAGCCAGTAGAATTGACTTGTGAGGCTCAACCCAAACAGCTTTTGGCAGGCGA 4563
Db 4315 GACGTGAGCTGACACAGATGTTGTGAGATGCAACCGGACAGCAGCTGTGTGGTCTG 4374
QY 4564 CCAAGTTATCTCCGATATGAGGCGAAGTGTATTCATACCTGGAAGGTACACGATTCAT 4623
Db 4375 AAGGCTTACAGTACCACTGACGGGTCGCTGTACTGTACTTTGAAGGTACGAAATTCAC 4434
QY 4624 CAGACCGCCAAAGACATTTGCCGAATCCATGATGTCGCCCAAAATCTGAGGCTAAT 4683
Db 4435 CAGGCTGCTATTGATATGTGCAAGATTAAGTACGTTGTGCCCAAGCTGCAAGAGCAAAAC 4494
QY 4684 GAGCAGATTGCTTGTATCATCTTGGGGAGATGTCCAGCATCCGCTCCAAATGCCCA 4743
Db 4495 GAAACAGATATGCTATATGCGCTGGGCGAAACAATGGAACAATCAGATCCAAATGTCCG 4554
QY 4744 GTAGAGAGTGAAGGCGTCTGCTCCACTCACACTTCCATGCTGTGTATTAATACGCT 4803
Db 4555 GTGAACGATTCGATTAATCAACCTCCAGACAGTGCCTGTGCTGCGCTACGCA 4614
QY 4804 ATGACGCTGAGCGCGTATACAGGTTGGCTCTGCGAAGAAAGAACAGTTGCCGTAATGC 4863
Db 4615 ATGACAGCAAGACCGATCGCCCGCTTAGTCAACACAGTTAAAGCATGTGTGTTGC 4674
QY 4864 TCATCATTTCTGTTGCCGAGTACAGATCAACAGCGGTGCAAGAACTACAGTGCAGCAA 4923
Db 4675 TCATCTTTTCCCTTCCGAAATACATGTAGATGGGGTGCAAGAGTAAAGTGCAGAGAG 4734
QY 4924 CCAATCTGTTTTCAGGCGTGTACCAACCGGCTGTACACCCCAAGAAATAGCCGGA 4980
Db 4735 GTTCTCTGTTGACCCGACGTAATTTCAGTGTGTAGTCCGCGAAGTATGCGCA 4791

Sindbis virus cDNA
Alphavirus-based e
Sindbis-like virus
Representative eu
Representative eu
Representative eu
Representative eu
Representative eu
RNA of expression
Sincchironp virus
Sindichiron virus
Polynucleotide se
Plasmid pRES-2J
Venezuelan equine
Venezuelan equine
Polynucleotide se
Plasmid pGDH-EB
Polynucleotide se
987Bneo DNA sequ
DSP6-SFV4 RNA tra
Venezuelan equine
Human immunodef
Venezuelan equine
Venezuelan equine
Sindbis virus cDNA
Sindbis virus vari
Sindbis virus vari
Sindbis virus cDNA
Polynucleotide sec
Sindbis expression
Alphaviral vector
Nucleotid sequenc

PT encephalitis virus genome - useful for the production of live or
PT attenuated vaccines for human or veterinary medicine

PS Example 1; Page 54-67; 112pp; English.

XX This CDNA sequence codes for an infectious western equine
CC encephalitis (WEE) virus RNA transcript. DNA representing the
CC entire genome was prepared by PCR using primers (see AA74110-21)
CC based on partial genome sequences. 5' Sequences were obtained by
CC RACE. The full-length infectious clone is useful in the production
CC of virulent WEE virus, and for introducing and testing attenuating
CC mutations. Also new are: infectious or attenuated WEE RNA
CC transcripts and WEE viral particles; cDNA (see AA74108) encoding an
CC infectious Venezuelan equine encephalitis (VEE) variant IE viral
CC genome; infectious or attenuated VEE RNA transcripts and viral
CC particles; an attenuated chimeric virus containing non-structural
CC sequences from a first alpha-virus (AV) and structural sequences
CC from a second AV, resulting in attenuation of the second AV; a method
CC for expressing a protein (especially an antigen to protect against
CC a pathogen) by cloning its gene into an attenuated WEE or VEE IE
CC replicon so that transcription of the replicon produces RNA able to
CC infect the cells in which protein is to be produced; methods for
CC diagnosing WEE and VEE infections; and polypeptides encoded by VEE
CC variant IIIA. Attenuated WEE and VEE are used in live or inactivated
CC vaccines, for use in human or veterinary medicine. Chimeric viruses
CC are also useful as vaccines, directed against the second AV which is
CC particularly Eastern equine encephalitis (EEE) virus or some variant
CC of VEE. WEE and VEE nucleic acids are used as primers and probes to
CC diagnose virus infections and to define natural variants, also for
CC production of protein antigens which can be used as diagnostic
CC reagents, to generate antibodies, and in vaccines. The attenuated
CC viruses are highly immunogenic and provide long-lasting protection.

XX Sequence 11492 BP; 3257 A; 2854 C; 2794 G; 2570 T; 17 other:

Query Match 90.6%; Score 5877.6; DB 20; Length 11492;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 6238; Conservative 0; Mismatches 221; Indels 29; Gaps 20;

QY 1 CACCGCCAGCAGCAACCGTAATATGTGAACCACTGTGCCAGAACGTATCCCACTC 60
DB 5012 CATGCCCAACCAACCAACCTTAATATGTGAACCACTGTGCCAGAACGTATCCCACTC 5071
QY CGGTGATTTCTAGAGCAACAGTGGGGAATCACTGCTATGCGGGGCGCTCGTTCT 120
DB 5072 CGGTGATTTCTAGAGCAACAGTGGGGAATCACTGCTATGCGGGGCGCTCGTTCT 5131
QY 121 CTAGCTGTGCCACAGCTGCTGCAACCGCTGGAGCGACTATAGACAGGCGTTGTGTTA 180
DB 5132 CTAGCTGTGCCACAGCTGCTGCAACCGCTGGAGCGACTATAGACAGGCGTTGTGTTA 5191
QY 181 CAGTGATGTGCATCAAGCGAACAGTCTACGTGAGAGCATCCCTAGTGTCTTGGCTTGG 240
DB 5192 CAGTGATGTGCATCAAGCGAACAGTCTACGTGAGAGCATCCCTAGTGTCTTGGCTTGG 5251
QY 241 ACGTCCAGCTGCGCTTGAAGCTCACTGATTCCTGAGATTCCTCAAGTGCACACAGGCT 300
DB 5252 ACGTCCAGCTGCGCTTGAAGCTCACTGATTCCTGAGATTCCTCAAGTGCACACAGGCT 5311
QY 301 TTGAAGTGAAGAACCACTCTGTACAGACCTTAACGTGCGAGTGTGCGAAGCTGTGGAC 360
DB 5312 TTGAAGTGAAGAACCACTCTGTACAGACCTTAACGTGCGAGTGTGCGAAGCTGTGGAC 5371
QY 361 TGGCCGAATATATCAAGACTTCAATACTGCTCTTTCAGTTTCTTTGGACTACAGAC 420
DB 5372 TGGCCGAATATATCAAGACTTCAATACTGCTCTTTCAGTTTCTTTGGACTACAGAC 5431
QY 421 CAGTACCGGACCAACGAGAGAGCCCATCCATCACTAGATGCAAGGCTTCCGCACTTC 480
DB 5432 CAGTACCGGACCAACGAGAGAGCCCATCCATCACTAGATGCAAGGCTTCCGCACTTC 5491
QY 481 CAGTCCAAAGCCAGCAGACTAAGTACCAACACCAAGAGTGCCTGAGACGATCT 540
DB CAGTCCAAAGCCAGCAGACTAAGTACCAACACCAAGAGTGCCTGAGACGATCT 5460

DB 5492 CAGTCCAAAGCCAGCAGACTAAGTACCAACACCAAGAGTGCCTGAGACGATCT 5551
QY 541 CAGTCCAAAGCCAGCAGACTAAGTACCAACACCAAGAGTGCCTGAGACGATCT 600
DB 5552 CAGTCCAAAGCCAGCAGACTAAGTACCAACACCAAGAGTGCCTGAGACGATCT 5610
QY 601 CAGTCCAAAGCCAGCAGACTAAGTACCAACACCAAGAGTGCCTGAGACGATCT 660
DB 5611 CAGTCCAAAGCCAGCAGACTAAGTACCAACACCAAGAGTGCCTGAGACGATCT 5667
QY 661 GTAAATACCAAGAACCTATATGTGATCGGCGCTGCAAGAGTATTAAGCCCGCGCC 720
DB 5668 GTAAATACCAAGAACCTATATGTGATCGGCGCTGCAAGAGTATTAAGCCCGCGCC 5726
QY 721 TCGATCTCGAAAGAGAGAAAATGTTACAGAAAGCTGCAATTAATGCGCTC-TGAAGGA 779
DB 5727 TCGATCTCGAAAGAGAGAAAATGTTACAGAAAGCTGCAATTAATGCGCTC 5786
QY 780 AATGAAAGCAGTATCAATACAGAAAGTA-GAAATATGAAAGCAATTAACAGCGAGCG 838
DB 5787 AATGAAAGCAGTATCAATACAGAAAGTA-GAAATATGAAAGCAATTAACAGCGAGCG 5846
QY 839 ACTCATTTCTGATTTGGGCAATATCATATCATGAGAGTGAATCTGTGAGTGTACAG 898
DB 5847 ACTCATTTCTGATTTGGGCAATATCATATCATGAGAGTGAATCTGTGAGTGTACAG 5906
QY 899 AGTCAATATTCCTGTACCAATCTACTGCTCAAGAGTGAATTAACAGTGTACATCTGAGA 958
DB 5907 AGTCAATATTCCTGTACCAATCTACTGCTCAAGAGTGAATTAACAGTGTACATCTGAGA 5966
QY 959 GGTGCGCGTTAAACGTCAGCACTTAAGTATCAAGAGAAATTAACCTTCAGTAGCCAGTTA 1018
DB 5967 AGTGGCGCTTAAAGTTTGCACTTATGATCTCAAGAGAAATTAACCTTCAGTAGCCAGTTA 6026
QY 1019 TTGATTAACAGATGAATACGATCGATCTGATCTGATGAGAGCGGCGCATCTGTCTCT 1078
DB 6027 TTGATTAACAGATGAATACGATCGATCTGATCTGATGAGAGCGGCGCATCTGTCTCT 6084
QY 1079 AGATACAGCACTTTTGTGCGGCTAACTGAGAAAGTACCAGAAAGAGTACGTTATTT 1138
DB 6085 AGATACAGCACTTTTGTGCGGCTAACTGAGAAAGTACCAGAAAGAGTACGTTATTT 6140
QY 1139 GCAGCCAGAGATTAAGATCAACCGTCCATCGCTATACAGAAATCAATTAACAAATGATT 1198
DB 6141 GCAGCCAGAGATTAAGATCAACCGTCCATCGCTATACAGAAATCAATTAACAAATGATT 6200
QY 1199 GCGTGCAGCTACTAAAGGAATGCAACGTTACCCCAATGCGAAGTATTAACCTGTCTTGA 1258
DB 6201 -GCTGCAGCCACTAAAGGAATGCAACGTTACCCCAATGCGAAGTATTAACCTGTCTTGA 6257
QY 1259 TTGGCGCGCATTTAATGTTTATGTTTCAAGAAATACGATGCAATGATAGTACGCGGA 1318
DB 6258 TTGGCGCGCATTTAATGTTTATGTTTCAAGAAATACGATGCAATGATAGTATGCGGA 6317
QY 1319 TACCTTGGCGATTAACCTTATGCGCTTAACCTACAGAGAAAGTTACGCAATGTGACAAA 1378
DB 6318 TACCTTGGCGATTAACCTTATGCGCTTAACCTACAGAGAAAGTTACGCAATGTGACAAA 6377
QY 1379 GCTGAAAGGCGCGAAGCAGACAGATTTGTCGAATATCTCATATTAATAACCGTTGCA 1438
DB 6378 GCTGAAAGGCGCGAAGCAGACAGATTTGTCGAATATCTCATATTAATAACCGTTGCA 6437
QY 1439 GGAATATCCATATGATTAATTCGTATGATCTTAAGAGAGATGTCAAAGTTACTCCCGG 1498
DB 6438 GGAATATCCATATGATTAATTCGTATGATCTTAAGAGAGATGTCAAAGTTACTCCCGG 6496
QY 1499 CAGAAACATATACAGAGAGCGGCTAAGTGAAGTATTAACAGCTGACAGTCCCTTGC 1558
DB 6497 CAGAAACATATACAGAGAGCGGCTAAGTGAAGTATTAACAGCTGACAGTCCCTTGC 6556
QY 1559 TACGCTTACCTTTGCGGAGTTCATCGGGAATTAAGTCCGTAGACTGATGCTTCT 1618
DB 6557 TACGCTTACCTTTGCGGAGTTCATCGGGAATTAAGTCCGTAGACTGATGCTTCT 6614

QY	1619	GCACAATATCCAACTCTCTTCGACATGTCAGCGGAGATTTTGATGCGATTAATGCTGA	1678
Db	6615	GCCGAATATCCAACTCTCTTCGACATGTCAGCGGAGATTTTGATGCGATTAATGCTGA	6674
QY	1679	ACATTTCCACACCGCGGACCCAGTATTTGGAAAACGACATCCGCTGTTTGAATTAAGGA	1738
Db	6675	ACATTTCCACACCGCGGACCCAGTATTTGGAAAACGACATCCGCTGTTTGAATTAAGGA	6734
QY	1739	AGACGACGCTATCCGCTATTTGCGGCGCTTGATGATCCTTGAGACCTTAGGTGACAAAC	1798
Db	6735	AGACGACGCTATCCGCTATTTGCGGCGCTTGATGATCCTTGAGACCTTAGGTGACAAAC	6794
QY	1799	GCTCTTAGATTTGATAGAGCGCGGCTTGCGCAATATCAATCTGTGCACTTACCTACAG	1858
Db	6795	GCTCTTAGATTTGATAGAGCGCGGCTTGCGCAATATCAATCTGTGCACTTACCTACAG	6854
QY	1859	AACGAGGTTTAAATTTGGTGCATGATGAAATCCGGTATGTTTCTTAACGCTGTTTGCA	1918
Db	6855	AACGAGGTTTAAATTTGGTGCATGATGAAATCCGGTATGTTTCTTAACGCTGTTTGCA	6914
QY	1919	CACACTAGTCAATATCATGATTTGCTAGAGAGTACTAGGTAAAGGGTTAAACAGTCAAC	1978
Db	6915	CACACTAGTCAATATCATGATTTGCTAGAGAGTACTAGGTAAAGGGTTAAACAGTCAAC	6974
QY	1979	GTCGCGGCGCTCTATTCGGCGGACGATTAACATAGTGATGTCGTCTCCACACTTGAT	2038
Db	6975	GTCGCGGCGCTCTATTCGGCGGACGATTAACATAGTGATGTCGTCTCCACACTTGAT	7032
QY	2039	GCGGAGAGATGCGGCCACTTTGGCTGACATGGAAATTAATTTATGTCAGTTATTTGG	2098
Db	7033	GCGGAGAGATGCGGCCACTTTGGCTGACATGGAAATTAATTTATGTCAGTCAATTTGG	7092
QY	2099	TATCAAAACACCCCTACTCTGTGGGGGGATTAATCCTGGTGAACAGATTAACAGCACAC	2158
Db	7093	TATCAAAACACCCCTACTCTGTGGGGGGATTAATCCTGGTGAATCAGATTAAC-GACACAC	7150
QY	2159	CTGCAGATGCGAGACCCCTCTTAAAGAGGCTTTTAAAGCTTTGGAAGAAACCATTTGCGAGT	2218
Db	7151	CTGTAG-GTCGAGACCCCTCTTAAAGAGGCTTTTAAAGCTTTGGAAGAAACCATTTGCGAGT	7209
QY	2219	TGATATCCCAAGACTGCGACCGCGCGCGGACCTGATGATGAAGCAATGCGATGGAAACAG	2278
Db	7210	TGACACCCCAAGACTGCGACCGCGCGCGGACCTGATGATGAAGCAATGCGATGGAAACAG	7269
QY	2279	AATTGGATTTACGAGACGAGTGAAGAAAGCCGTGAATCCAGATTAACGAGATCATATCGGC	2338
Db	7270	AATTGGATTTACGAGACGAGTGAAGAAAGCCGTGAATCCAGATTAACGAGATCATATCGGC	7329
QY	2339	AGGCTGATCATACAGTCTCTGTCCAGCTTAAGCCGAAGCGTTAAGAACTTCAAGAGCAT	2398
Db	7330	AGGCTGATCATACAGTCTCTGTCCAGCTTAAGCCGAAGCGTTAAGAACTTCAAGAGCAT	7389
QY	2399	AAGAGGAGCCCAATCACCTCTTACGCGTGAACCTAAATAGGTGAAGTATGAGACACGAC	2458
Db	7390	AAGAGGAGCCCAATCACCTCTTACGCGTGAACCTAAATAGGTGAAGTATGAGACACGAC	7449
QY	2459	CTACCCACCGGC-AGAATGTTTCCATACCTCTCAGCTGAACCTTTCCACAGTTTACCTTAC	2517
Db	7450	CTACCCACCGGC-AGAATGTTTCCATACCTCTCAGCTGAACCTTTCCACAGTTTACCTTAC	7509
QY	2518	AAATCCGATGGCTTACGAGATTCGAAACCCCTCTAGGCGCGGCTGGAGACCGTTTCGGC	2577
Db	7510	AAATCCGATGGCTTACGAGATTCGAAACCCCTCTCTTAGGCGCGGCTGGAGACCGTTTCGGC	7569
QY	2578	CCCGCTGGCTGCTCAATTCGAAGATCTTAGAGAGTGCATATGATCACTTGACTTTCAACA	2637
Db	7570	CCCGCTGGCTGCTCAATTCGAAGATCTTAGAGAGTGCATATGATCACTTGACTTTCAACA	7629
QY	2638	ACGATACCTTAATCCGCGCGCCAGGTGCACCGCAAGAAAGAAAGAGTCTCTTAAGCC	2697
Db	7630	ACGATACCTTAATCCGCGCGCCAGGTGCACCGCGCAAGAAAGAAAGAGTCTCTTAAGCC	7689

QY	2658	AAAACTTACTACGCTTAAAAAGAAAGACAGCAACCCAAAGGACGAAAGCCGACCTTAA	275
Db	7690	AAACCTTACTAGCTTAAAAAGAAAGACAGCAACCCAAAGGACGAAAGCCGACCTTAA	7749
QY	2758	ACCAAGGAAAGCAACGATATGTGTATGAATTTGAGAGTCGGACAAAGACATTTCCGATCAT	2817
Db	7750	ACCAAGGAAAGCAACGCTATGTGTATGAATTTGAGAGTCGGACAAAGACATTTCCGATCAT	7808
QY	2818	GCTGAACGGCCAAAGTGAATGATATGCTTCGCTTGTCCGAGGAAGCTGATGAACAACCT	2877
Db	7810	GTTGAACGGCCAAAGTGAATGATATGCTTCGCTTGTCCGAGGAAGGCTGATGAACAACCT	7865
QY	2878	CCACGTTGAAGAAAAAATTGATTAATGACCAATTACCGCCGCTGAATTGAAGAAAGCTAG	2937
Db	7870	CCACGTTGAAGAAAAAATTGATTAATGACCAATTAGCCGCTGAATTGAAGAAAGCTAG	7929
QY	2938	CATGTAGCACTTGGAGTACCGGACGCTTCCCGACAAACATGAAATTCAGACAGCTGCAAGTA	2997
Db	7930	CATGTAGCACTTGGAGTATGGCGACGCTTCCCGAATATGAAATTCAGACAGCTGCAAGTA	7988
QY	2998	CACCAAGCACAACCAACCGGGCTTCTACATCTGGCACCACGGCCGACGTCAGTATGAGAA	3057
Db	7990	CACCAAGCACAACCAACCGGGCTTCTACATCTGGCACCACGGCCGACGTCAGTATGAGAA	8049
QY	3058	TGGAGATTTTACCGTACCAGAGAGATGGGCGGAAAGCGACAGCCGAGAACCCGATCT	3117
Db	8050	TGGAGATTTTACCGTACCAGAGAGATGGGCGGAAAGCGACAGTGGAAAGCCGATCTCT	8109
QY	3118	GGAACAACGAGGACAGATGTGTGCTATTGTCTTAGAGATCCAAATGAGGCGACGGCTAC	3177
Db	8110	GGAACAACGAGGACAGAGTGTGTGCTATTGTCTTAGAGAGTCCAAACGAGGCGACGGCTAC	8168
QY	3178	GAGCGTTTCAGTGTGCTACTTGGAAACCAAGAAAGGGGTGACATTAAGGATACCCCGAAGG	3237
Db	8170	GAGCGTTTCAGTGTGCTACTTGGAAACCAAGAAAGGGGTGACATTAAGGATACCCCGAAGG	8229
QY	3238	TTCTGAACCGTGTGCTACTTAGTAAACGCGCTATGCGTGTCTTTCGAAATGTCACGTTCCATG	3297
Db	8230	TTCTGAACCGTGTGCTACTTAGTAAACGCGCTATGCGTGTCTTTCGAAATGTCATTTCCCTTG	8289
QY	3298	CGACAAACCAACCCGTGTGCTATTCACTGACCCCAAGAACATCTCCACGTGTCTGAAAGA	3357
Db	8290	CGACAAACCAACCCGTGTGCTATTCACTGACCCCAAGAACCACTCCGACGTGTCTGAAAGA	8349
QY	3358	GAACTGTGACATCCAAATTACGACACGCTGCTGAGAAAGCTCTTGAATAATGTCATCACG	3417
Db	8350	GAACTGTGACATCCAAATTACGACACGCTGCTGAGAAAGCTCTTGAATAATGTCATCACG	8408
QY	3418	CCGGCCCAACGAAGCATTAACGATGACTTCACACTGACCAAGTCCCTTACTTGGGGGTTCTG	3477
Db	8410	CCGGCCCAACGAAGCATTAACGATGACTTCACGACTGACCAAGTCCCTTACTTGGGGGTTCTG	8466
QY	3478	CCCGTATTGCAACACTCAACGCGGTGTTTACGCCCAATAAAAATTGAAACGTTGSGGA	3537
Db	8470	CCCGTATTGCAACACTCAACGCGCATGTTTACGCCCAATAAAAATTGAAACGTTGSGGA	8529
QY	3538	CGAATCTGATATGAGTGAATTAACATTCACAGTCTCCGGCAAAATTCGGCTTCAATCAAGC	3597
Db	8530	CGAATCTGATATGAGTGAATTAACATTCACAGTCTCCGGCAAAATTCGGCTTCAATCAAGC	8589
QY	3598	AGGCACTGCGGATGTCAACCAAAATTCGGTATCATGCTTTGACACGACGCACTGACATCAA	3657
Db	8590	AGGCACTGCGACGTCACCAAGTTCCGGTACATGCTTTACGACCAACGACATGACATCAA	8649
QY	3658	GGAAAGACGATATGAGAAAAATAGCTATCAGACATCTGGAACCTTCGGCTGCTTTGCGCA	3717
Db	8650	GGAAAGACGATATGAGAAAAATAGCTATTAATGATCACTGGACCAATGCGCTCTTTCGCGCA	8709
QY	3718	CAAAAGGATCTTCCGTTAGTCAATGCTCAATGCTTCACAGGTGACAGTGTAAACCCGTCAGATATCAC	3777
Db	8710	CAAAAGGATCTTCCGTTAGTCAATGCTTCACAGGTGACAGTGTAAACCCGTCAGATATCAC	8769
QY	3778	GAGCGAGCATCTGTAGATTCATCACCGTGTGAGAAAAAGATTCAGAGAAAGTTGTTCG	3837

|||||
Db 8770 GAGGGAGCATCTGAGATTCATGACCGGTGAGAAAAAGATCAGAGGAAGTTTGTCCG 8829
Qy 3838 TAGAGAGAGATCTTTTCCACCCGTCATGAGAAAGCTGTAAAGTCCACGTTTACGA 3897
Db 8830 TAGAGAGAGATCTTTTCCACCGTCATGAGAAAGCTGTAAAGTCCACGTTTACGA 8889
Qy 3898 TCATCTGAGAGAGAGCTCTGCGGGTACATACCATGACAGGCGAGGCCACAGCGTA 3957
Db 8890 TCATCTGAGAGAGAGCTCTGCGGGTACATACCATGACAGGCGAGGCCACAGCGTA 8949
Qy 3958 TAACTCTATCTGAGAGAGCGTCAGGGAAGTGTACATTAAACACCTTCTGCAAGAA 4017
Db 8950 TAACTCTATCTGAGAGAGCGTCAGGGAAGTGTACATTAAACACCTTCTGCAAGAA 9009
Qy 4018 CGTCACTACGATGATGAGTGTGCGGACTACAGCAGGTAATCGTAGACGCGAAGAA 4077
Db 9010 CGTCACTACGATGATGAGTGTGCGGACTACAGCAGGTAATCGTAGACGCGAAGAA 9069
Qy 4078 GATGAAGGGCTGCACTAAGCAAAAGTGATTCGCTACAGAGCGACCAAGGAATG 4137
Db 9070 GATGAAGGGCTGCACTAAGCAAAAGTGATTCGCTACAGAGCGACCAAGGAATG 9129
Qy 4138 GGTCTCAACTCGCGGATCTTATTAGGCACACAGACCACTCAGTGCAGAGGTAATTGCA 4197
Db 9130 GGTCTCAACTCGCGGATCTTATTAGGCACACAGACCACTCAGTGCAGAGGTAATTGCA 9189
Qy 4198 CATTCATTCGCGCTTGACACCGACAGCTGCGCGGTTCCGTTAGTCTACAGGCTACAGT 4257
Db 9190 CATTCATTCGCGCTTGACACCGACAGCTGCGCGGTTCCGTTAGTCTACAGGCTACAGT 9249
Qy 4258 CACGAAGGTTCAAAGGATCACCCCTCACCTGACTGCAATGAGCAACCAATTCGTGAC 4317
Db 9250 CACGAAGGTTCAAAGGATCACCCCTCACCTGACTGCAATGAGCAACCAATTCGTGAC 9309
Qy 4318 AACGAGAAATTTGGGCTGCGAGAGAGCAAGCAAGCAGAAATGATTAAGGCTTACATC 4377
Db 9310 AACGAGAAATTTGGGCTGCGAGAGAGCAAGCAAGCAGAAATGATTAAGGAGCTACATC 9369
Qy 4378 CAGAAATTTTCTGTGGGGCCAGAAAGGCTGAGTACGTATGGGGTAACTGAACCACT 4437
Db 9370 CAGAAATTTTCTGTGGGGCCAGAAAGGCTGAGTACGTATGGGGTAACTGAACCACT 9429
Qy 4438 CAGAGTCTGGGCCAGAGAGTGGGACCAAGGCAACCAATGATGGCGCATGAGATCAT 4497
Db 9430 CAGAGTCTGGGCCAGAGAGTGGGACCAAGGCAACCAATGATGGCGCATGAGATCAT 9489
Qy 4498 CATCACTATTATCATCGGCATCCAGTCTACACTGTCTGCTGTGTGTGCTCT 4557
Db 9490 CATCACTATTATCATCGGCATCCAGTCTACACTGTCTGCTGTGTGTGCTCT 9549
Qy 4558 TGCATCTCTGTAGGACATGATCATGACGAGCTTGATCGCCAAAGMAAGAGACTG 4617
Db 9550 TGCATCTCTGTAGGACATGATCATGACGAGCTTGATCGCCAAAGMAAGAGACTG 9609
Qy 4618 CCTGACGCATACGCGCTTGACCGAAGCAACGGTACCAAGATTAAGGGGTTTGTG 4677
Db 9610 CCTGACGCATACGCGCTTGACCGAAGCAACGGTACCAAGATTAAGGGGTTTGTG 9669
Qy 4678 CTGATTTGGGCCAACAACGCTGAACAATTTGAGAAAATTGAAACCATCTGTGGTTAA 4737
Db 9670 CTGATTTGGGCCAACAACGCTGAACAATTTGAGAAAATTGAAACCATCTGTGGTTAA 9729
Qy 4738 CAACCAACCGTTTCTCTGGGCAAGTGTGATTCCTCTGGCAGCGCTTGTATTCTGTT 4797
Db 9730 CAACCAACCGTTTCTCTGGGCAAGTGTGATTCCTCTGGCAGCGCTTGTATTCTGTT 9789
Qy 4798 CCGGCTCTTTCATGCTGCATGCTTTTATTGGTGCAGGCGCTGCGGGAAAGT 4857
Db 9790 CCGGCTCTTTCATGCTGCATGCTTTTATTGGTGCAGGCGCTGCGGGAAAGT 9849
Qy 4858 AGAGCGCTTTCGAACATGCGACACTGTGCAAAATGTTCCGGGAGTCCCGTATAAGGCGCTT 4917
|||||

Db 9850 AGAGCGCTTTCGAACATGCGACCACTGTGCCAAATGTTCCGGGATCCCGTATAAGCGCTT 9909
Qy 4918 GGTGAAACGGCGAGTTAGCGGCACTTAACCTGAGATCAGCGGTGCTCATCGGAATT 4977
Db 9910 GGTGAAACGGCGAGTTAGCGGCACTTAACCTGAGATCAGCGGTGCTCATCGGAATT 9969
Qy 4978 AACACCTTCAACTAACAGAGATACGTGACCTGCMAATTCACAGATCATTCCTTACC 5037
Db 9970 AACACCTTCAACTAACAGAGATACGTGACCTGCMAATTCACAGATCATTCCTTACC 10029
Qy 5038 ACAAGTTAAATGCTGGGGTCCCTCGAGTGCAGAGCATCTCTCAAGCGGATTAACATG 5097
Db 10030 ACAAGTTAAATGCTGGGGTCCCTCGAGTGCAGAGCATCTCTCAAGCGGATTAACATG 10089
Qy 5098 CCGGCTTTTGGGGTGTGTAACCTTTCATGTTGGGGGGGCAATGCTTCTGTGACAG 5157
Db 10090 CCGGCTTTTGGGGTGTGTAACCTTTCATGTTGGGGGGGCAACAGTCTTCTGTGACAG 10149
Qy 5158 TGAGAACACAACTGAGTGAAGGCGTACGTCGAGTTCGCTCCAGACTGCATATAGATCA 5217
Db 10150 TGAGAACACAACTGAGTGAAGGCGTACGTCGAGTTCGCTCCAGACTGCATATAGATCA 10209
Qy 5218 CGCAGTCGACTTAAAGTTCAACAGCTGCTTGAAAGTGGCCCTGCTATAGTATACG 5277
Db 10210 TGCAGTCGACTTAAAGTTCAACAGCTGCTTGAAAGTGGCCCTGCTATAGTATACG 10269
Qy 5278 CAACACACCGCGCACTGATAGCTTACGTTGTCATAGGCGTACGCGAGGTTCTTACAGGGA 5337
Db 10270 CAATACCAACGCGGCTGATACATTCGTACACGCGGTCAACACAGGTTCTTACAGGGA 10329
Qy 5338 CTTGAAGTCAATAGCAGGGCGATATCAGCGGCTTTTCAACCTTTGACCATTAAGTCTG 5397
Db 10330 CTTGAAGTCAATAGCAGGGCGATATCAGCAGGCTTTTCAACCTTTGACCATTAAGTCTG 10389
Qy 5398 CATCAGAAAGGCGCTTCTTCAACTACGACTTCCTGAGTATGAGACTATGAACCAAG 5457
Db 10390 CATCAGAAAGGCGGCTTCTTCAACTACGACTTCCTGAGTATGAGACTATGAACCAAG 10449
Qy 5458 AGGTTGGGGATTAATGAACATCTCGCTGATGTCAGACATAGTAGCCGCACTGA 5517
Db 10450 AGGTTGGGGATTAATGAACATCTCGCTGATGTCAGACATAGTAGCCGCACTGA 10509
Qy 5518 CATACGGCTCTGAAGCTTCTGTCAAGAACATCCAGTCCCTTACCCCAAGCAGTATC 5577
Db 10510 CATACGGCTCTGAAGCTTCTGTCAAGAACATCCAGTCCCTTACCCCAAGCAGTATC 10569
Qy 5578 AGGTTATGAATGTGAAGAACAACTCAGACGACCCCTCAGAAACACAGCACTTTGG 5637
Db 10570 AGGTTATGAATGTGAAGAACAACTCAGACGACCCCTCAGAAACACAGCACTTTGG 10629
Qy 5638 ATGTAAATTTGAAGTGAAGCTCGAGGCTCTAACCTGTATGAGGGGACATCCCTAT 5697
Db 10630 ATGTAAATTTGAAGTGAAGCTCGAGGCTCTAACCTGTATGAGGGGACATCCCTAT 10689
Qy 5698 CTCGATTCAGATCCCTGATGACGCTTGTGAGATCATCAGAAATCAACCAATTTTGA 5757
Db 10690 CTCGATTCAGATCCCTGATGACGCTTGTGAGATCATCAGAAATCAACCAATTTTGA 10749
Qy 5758 AGTTAGCTGACAGTACAGACTGATTTATTTGCGAGACTTTGGTGTCTCTTAACTT 5817
Db 10750 AGTTAGCTGACAGTACAGACTGATTTATTTGCGAGACTTTGGTGTGTCTCTTAACTT 10809
Qy 5818 ACAGTACAAAGCTGACAGGGAGGGAATGTGCCAGTTCATCCCACTCCAGACAGCTGT 5877
Db 10810 ACAGTACAAAGCTGACAGGGAGGGAATGTGCCAGTTCATCCCACTCCAGACAGCTGT 10869
Qy 5878 TTTGAGAGAGCGACCAACATGATGATGCGGTAGGAGCATTAACATTTTACAC 5937
Db 10870 TTTGAGAGAGCGACCAACATGATGATGCGGTAGGAGCATTAACATTTTACAC 10926
Qy 5938 ATCAGGCCCAACAGCAAAATTTATAGTTTGCTATGCGGCAAGAGTCCACTGCAATGC 5997
Db 10927 ATCAGGCCCAACAGCAAAATTTATAGTTTGCTATGCGGCAAGAGACCACTGCAATGC 10986
|||||

```

QY 5998 TGAATGTAACACCGCGGACACATATTTGGAGAACACATAAAGTCGACCAAGAAAT 6057
DB 10987 TGAATGTAACACCGCGGACACATATTTGGAGAACACATAAAGTCGACCAAGAAAT 11046
QY 6058 CCAGCGCGGAGTTTCCAAAACATCTTGGAACTGGCTGTGACGTTTGGGGAGCATC 6117
DB 11047 CCAGCGCGGAGTTTCCAAAACATCTTGGAACTGGCTGTGACGTTTGGGGAGCATC 11106
QY 6118 ATCCCTCATGTTGTTAGAGCTATATGTTGTTGCTGACGCTTATGCTTATAACACACG 6177
DB 11107 ATCCCTCATGTTGTTAGAGCTATATGTTGTTGCTGACGCTTATGCTTATAACACACG 11166
QY 6178 TAGATGACGACGCGGACACTGACATAGCGGT-AAAAGTCGATGACTCCGAGGAAGC 6236
DB 11167 TAGATGACGACGCGGACACTGACATAGCGGTAAAAGTCGATGACTCCGAGGAAGC 11226
QY 6237 GTGGTGCAATATGCCACGCGCGCTTGACACTAAAAGTCGATGATTTCCGAGGAAGC 6296
DB 11227 GTGGTGCAATATGCCACGCGCGCTTGACACTAAAAGTCGATGATTTCCGAGGAAGC 11286
QY 6297 AGTGATATAGCTGTGACAGTGTCACTATATGCTATATACACTACATTTTAAACACT 6356
DB 11287 AGTGATATAGCTGTGACAGTGTCACTATATGCTATATACACTACATTTTAAACACT 11346
QY 6357 ATATCACTTTTATGAGACTCACTATGCGTCTTAATATACACTACATTTTAA 6416
DB 11347 ATATCACTTTTATGAGACTCACTATGCGTCTTAATATACACTACATTTTAA 11406
QY 6417 AAACACTATACACTTTTATAAATCTTTTATTAATTTTCTTTGTTTATTTGTTT 6476
DB 11407 AAACACTATACACTTTTATAAATCTTTTATTAATTTTCTTTGTTTATTTGTTT 11466
QY 6477 TAAATTT 6484
DB 11467 TTAATAAT 11474

RESULT 2
AAV33475
ID AAV33475 standard; CDNA, 11663 BP.
AC AAV33475;
XX 10-DEC-1998 (first entry)
DT
DE South African Arbovirus strain No. 86 cDNA clone p555 sequence.
XX
XX Bone marrow cell; alphavirus; South African Arbovirus strain No. 86;
KW S.A.AR86; reverse transcriptase-PCR; nsP1; nsP2; nsP3; nsP4; capsid;
KW E3; E2; 6K; E1; growth hormone; growth factor; interleukin; cytokine;
KW chemokine; enzyme; ribozyme; antisense oligonucleotide; ss.
XX
OS South African arbovirus.
XX
XX Key Location/Qualifiers
XX FH 215
XX FT /*tag= a
XX FT /note= "This nucleotide varies from A to G with respect
XX FT variation to the S.A.AR86 cDNA sequence given in AAV33473"
XX FT 3863
XX FT /*tag= b
XX FT /note= "This nucleotide varies from C to G with respect
XX FT variation to the S.A.AR86 cDNA sequence given in AAV33473"
XX FT 5984
XX FT /*tag= c
XX FT /note= "This nucleotide varies from A to G with respect
XX FT variation to the S.A.AR86 cDNA sequence given in AAV33473"
XX FT 9113
XX FT /*tag= d
XX FT /note= "This nucleotide varies from T to C with respect
XX FT to the S.A.AR86 cDNA sequence given in AAV33473"
XX

```

```

PN WO9836779-A2.
XX 27-AUG-1998.
PD
XX 18-FEB-1998; 98WO-US02945.
PF
XX 19-FEB-1997; 97US-0801263.
PR
XX (UNNC-) UNIV NORTH CAROLINA.
PA
XX Davis NL, Johnston RE, Simpson DA,
XX WPI, 1998-495361/42.
DR
XX P-PDB; AAW70471, AAW70472.
DR
XX
PT Expression of heterologous RNA in bone marrow cells - using a
PT recombinant alphavirus comprising a promoter operable in bone marrow
PT cells operably associated with heterologous RNA.
XX
XX Example 5; Fig 5A-5B; 68pp; English.
XX
XX The present sequence represents the South African Arbovirus strain
XX No. 86 (S.A.AR86) cDNA clone p555 sequence used in the method of
XX the invention. The invention provides a method for introducing and
XX expressing heterologous RNA in bone marrow cells using alphavirus
XX vectors. S.A.AR86 is an alphavirus whose cDNA sequence (AAV33474) was
XX determined from uncloned reverse transcriptase-PCR reaction fragments
XX amplified from the virion RNA. The S.A.AR86 cDNA clone p555 sequence
XX has four nucleotide variations when compared with the S.A.AR86 cDNA
XX sequence derived from genomic RNA. The variations, however, do not
XX alter the sequence of the protein encoded with respect to the
XX S.A.AR86 genomic RNA. The S.A.AR86 cDNA encodes nonstructural
XX and structural polypeptides. The nonstructural polypeptide is,
XX presumably, post-translationally modified into four different protein
XX products, namely nsP1 (AAW70460), nsP2 (AAW70461), nsP3 (AAW70466), nsP4
XX (AAW70467) proteins. The structural polypeptide is, presumably,
XX post-translationally modified into five different protein products,
XX namely capsid (AAW70468), E3 (AAW70469), E2 (AAW70470), 6K (AAW70471),
XX E1 (AAW70472) proteins. The S.A.AR86 cDNA was used in the method of the
XX invention. The inventors claim the transformed bone marrow cells are
XX useful for expressing a protein or peptide suitable for protecting the
XX subject against a disease such as a microbial, bacterial, protozoal,
XX parasitic or viral disease. The transformed bone marrow cells are also
XX claimed to be useful for expressing proteins and peptides such as
XX hormones, growth hormones, growth factors, interleukins, cytokines,
XX chemokines, enzymes, ribozymes or antisense oligonucleotides.
XX
XX Sequence 11663 BP; 3289 A; 3105 C; 2910 G; 2359 T; 0 other;
SQ
XX
XX Query Match 37.0%; Score 2397.6; DB 19; Length 11663;
XX Best Local Similarity 64.6%; Pred. No. 0;
XX Matches 3752; Conservative 0; Mismatches 1984; Indels 74; Gaps 9;
XX
QY 580 ACTGACGATGAGACGCGGACGATATTTTCTCATCGGAAGACCGCAAGTCACCTTC 639
DB 5707 ACTGCTTAACCGGGGTAGTGTTGATCATTTTGGACGACACAGCCCTGGCACCCTTC 5766
QY 640 AACGAATTCAGTACGTCATGTAACTACAGAACTTATTTGATCGGACCTTCATG 699
DB 5767 AAAAGAGTCCGTTCTCAGAACGACGTTACAGAACGACCTTGGAGGCAATGTTCTGG 5826
QY 700 AGAAGTATTAAGCCCCCGCCCTCATCTCGAAAGAGGAAATGTACAGAAAGAACG 759
DB 5827 AAAAGATCTAAGCCCGGTGCTGACAGCTGAAAGAGGAAACACTCAACTCAGTACC 5886
QY 760 AATTATGCGCTCTGAGGAATATGAGCAGGTATCAATACGAAAGTAGAATAATATGA 819
DB 5887 AGATGATGCCACCGAAGCCCAACAAAGCAGTACGATCTCGAAAGTAGAATAACGGA 5946
QY 820 AAGCAATTACGCGGAGGACCTATTTCTGATTTGGGCAATATCTATCATCAGAACTGA 879
DB 5947 AAGCAATACCACTGAGGAGACTGCTTTCAGGAGCTACGAGCTGTATTAATCTTCCACAG--- 6003

```

QY	880	ATCTCTGTCGAGTGTTCACGAGTCAATTATCTGTATCCCAATCTACTCTGCTCAACGGTAATTA	939
Db	6004	ATCAGCCAGAAATGCTATTAAGATCACTTACCCGMAACCACTGATTTCCAGCAGTATACCAAG	6063
QY	940	ACAGGTTTACATCTTCCAGAGGTCGGGTAAACCGTCGACCTTAGTATCCAGAGAAATT	999
Db	6064	CGAACTACTCTGACCCCAAAAGTTTGCTGTAGCTGTTTGTAAACATATCTGATGAGAAATT	6123
QY	1000	ACCTTACAGTACGCGAGTTATGTATTAACAGATGAATACGATGCGTATCTTGACATGTGTGG	1059
Db	6124	ACCCGACGCTGAGACTCTTATCATGAGATCACCGACGAGTACGATGCTTACTTGGATATGTAG	6183
QY	1060	ACGGGCACTCGTCTGTCTATGATACAGCACTTTTGTCCGGCTTAACTGAGAACTATCC	1119
Db	6184	ACGGGACAGTCGCTTGCTCTAATATCTGMACTTTTGTCCCGCCAAAGTTTGAAGTTAAC	6243
QY	1120	CAAGAACATATAGCTATTTGGACGACAGATTAAGATTCAGCCGCTCCATCGCTTATACGA	1179
Db	6244	CGAAAAGACAGAGTATAGAGCCCAACATCCGATGTCGGTTTCATCAGCATGCAGA	6303
QY	1180	ATACATTTACAAAATGTATTTGGCTGCGAGCTACTTAAAGAAATTCGAACTGTACCCAAATGC	1239
Db	6304	ACAGGTTCCAAAACGTGCTCATTTGCGGACCTTAAAGAAACCTGCAACGTCAACAAATGC	6363
QY	1240	GAGAAATTACCTGTCTTAAATTCGCGCGCACTTAAATGTGTTGTTCAAGAAATTCGAT	1299
Db	6364	GTGAATCTCCCAACACTGCACTCAGCGAATTCACAGTTGAATGCTTTCGAAAATATGAT	6423
QY	1300	GCAATGATGAGTACTGGGATACCTTTGGCGAATACCTTATCGGCTAATCTACAGGAACG	1359
Db	6424	GCAATGACGAGTATTTGGAGGAGTTTGCCCGAAGCCATTTAGATCTACTGATGTTGC	6483
QY	1360	TTACGCAATATGTGACMAAGCTGMAAGGCGCGMAAGCAGCAGCAATTGTTTCGAATATCTC	1419
Db	6484	TTACCGCATACGTGGCCAGACCTGAAGGCGCTTAAGGCGCGCACTGTGTCCAAAGAACGC	6543
QY	1420	ATTAATCTTAAACCGTTGACAGAGATACCAATGATCAATTTGTCATGATCTTAAAGAG	1479
Db	6544	ATAATTTGTCCTCCATTTGCAAAAGTCCCTATGAGTATGATTCGTCACTGACATGTAAAGAG	6603
QY	1480	ATGTCAAAAGTTACTCCGGGACGAACATACAGAGGACCGGCTTAAGGTGAGGTTATTC	1539
Db	6604	ACGTGAAGTTTACACTTGCGACGAAACACACGAAAGAAAGCCGAATATCAAGTGATAC	6663
QY	1540	AGGCTGCAGATCCCTTTCCTACCGCTTACCTTTGGCGGATCATCGGAATTAATGCCGTA	1599
Db	6664	AAAGCGGAGAAACCCCTGGCGAACCGTTTACTATGCGGATCCACCGGGAGTTAATGTGCGCA	6723
QY	1600	GACTGAAATGCGGTGCTTTTGCCMAATATCATCTCTTTGCACATGTCCAGCGGAGATT	1659
Db	6724	GAGTTTACAGCGGTTTTGCTTACCCAATTTCAACGCTCTTTGACATGTCCGCGGAGGACT	6783
QY	1660	TTGATGCAATTTATGCTGTAACATTTTCCACACGCGGACCCGATATGGAACCGACATCG	1719
Db	6784	TTGATGCAATATAGCAAAACCTTCAAGACAAAGTACCCGTTACTGGAGACGATATACG	6843
QY	1720	CGTCGTTTGAATAAAGCGAAGACGACGTATCGCCATTTCCGCGTTGATGATCTCTTGAG	1779
Db	6844	CCTCGTTTGAACAAAAGCCAAAGACGCTATATGCGCTTAAACCGGCTGATGATCTTTGGAAG	6903
QY	1780	ACTTAGGTGTGCAGCAACCGCTCTTGAATTTGATAGAGCGGCGTTCCGCAATATCAT	1839
Db	6904	ACCTGGGTGTGACCAACACTACTGACTGATGAGTGGCGCTTTGGAAATATATCAT	6963
QY	1840	CTGTGCACTTACCTTACAGAAACGAGGTTAAATTTGTGTCATGATGAATTCGGTATGT	1899
Db	6964	CCACCCATCTGCCCAACGGGTACCCGTTTCAAAATTCGGGCGCATGATGAAATCCCGAATGT	7023
QY	1900	TCTTAAAGCTGTTTGTCAACACACTAGTCAATATCATATGATGTTGTAAGCAGATACATAG	1959
Db	7024	TCTTACGCTCTTTGTCAACACAGTTTCTGAATGTGTATATGCGCACGAGATATTTGAGG	7083
QY	1960	AACGGTTTACACGTCAAGCTGTCGGCGGCTCTATCGCGAGCATTAACATAGTGCATGGTG	2019

Db	7084	AGCGCTTAAACGTCMAATGTGTAGATTTATCGGAGACGACAACTTATTACCGAG	7143
Oy	2020	TGCTGTCCGACACCTTGATGCGGAGATGCGGCACCTTGCTGTAACATGGAATMAAA	2079
Db	7144	TAGTATCTGAAGAAAGAAATGGCTAGAGGTGTGGCACTGGCTCAATGGAGGTTAAG	7203
Oy	2080	TTATTGATGACGTTATTGTGTATCAAAGCACCTTACTTGTGGGGGATTTATCTGGTGG	2139
Db	7204	TCATTGACGCGAGTATCGGCGAGAGACCACTTACTTCTGGGCGGATTCATCTTTCGAA	7263
Oy	2140	ACCATTAACAGGACAGCGCTGCGAGAGTGGCAGACCCCTTAAAGAGCTTTTAAAGCTTG	2199
Db	7264	ATTGCGTTACCTCCACAGCGTGTGGCGTGGGAGCCCTTGAAGAGCTGTTTAAAGTTGG	7323
Oy	2200	GAAGAACATTCGACAGTCGATGATACCCAAAGCTCGACCGCGCGCGGACATGATATG	2259
Db	7324	GTAAGACCGCTCCGACGCGAGATAGGACAGACGAAGACGAAGACGGGCTCTGGCTAAGTG	7383
Oy	2260	AAGCAATGCGATGGAACGAATTTGGAATTACGACGAGTTGTGAAGGCGGTGAATTC	2319
Db	7384	AAACAAAGCGCTGGTTTGAAGTAGGTATACAGACACTTGTAGCATGGCCGTGCAATC	7443
Oy	2320	GATACGAGATCATACTGCGCAGCGCTGATCATCAGTCTCTGTCCACGTTAGCCGAAGCG	2379
Db	7444	GGTATGAGGTAGACAAACATCACACTGTGCTGTGGCATTTGAGAACTTTGGCCAGAGCA	7503
Oy	2380	TTAAGAACTTCAAGAGCATTAAGAGGAGGCCAATACCCCTTACGCGTGACCTTAAATAG	2439
Db	7504	AAAGGCACTTTCMAAGCCATCAGAGGGGAAATTAAGCATCTCTAGGTGTCTTAAATAGT	7563
Oy	2440	TGACGTAGTAGA-----CAGCACTTACCCACCGCAGATTTTCCA	2482
Db	7564	CAGCATAGTACATTTACTCTACTAATACCAACACACACCAACCATGATATGAGAGATTCT	7623
Oy	2483	TACCTCAGCTGAACCTTTCACAGATTATCCCTTCAAAATCCGATGCGTTACCGAGATCCA	2542
Db	7624	TTAATATGCTGGCGCGCCGCTTCCAGCCCCACCTGCTATGTGAAGCGCGGAGAA	7683
Oy	2543	AAACCTCTTAAGCGCGCGCTGAGGCGCGTTTGCGCCCGCTGCGTGTCAATCGAAGAT	2602
Db	7684	GGAGGACAGCGGCGCCCATGCTGCTGCGGCAATGGGCTGGCTTCCCAATCCAGCAATCG	7743
Oy	2603	CTTAGAGAGTGCATAGTCACTTGAACCTTTCAAACAC--GATCACTTAATTCGCGCGCAG	2660
Db	7744	CCACAGCGGTAGTGCCCTAGTCTCATTTGGACAGGCAACTAGACCTTCAAAACCCACGCCAC	7803
Oy	2661	GTTCAACCGCCCAAGAAAGAAAGAGTGTCTTAAGCCAAACCTAATCAGCTTAAAGAA	2720
Db	7804	GCCGCGCGCGCGCCAGAAAGAGGCGCCAAAGCAACCCGAAAGCCGGAAGAAACCA	7863
Oy	2721	AGAAAGCAACAGCCAAAGAGAGCAAGAAC--GCAAGCTTAAACAGAGGAAACGAACAAGTA	2777
Db	7864	AAACACAGGAAAGAAAGAAAGAACGACACTGCAAAACCCAAACCCGGAAGAGACAGCTA	7923
Oy	2835	ATGATATATGCTGCGGTTGTGCGAGAAAGCTGATGAAACCACTCAGCTTGAAGGAAAA	2894
Db	7984	TCGGGCAACGCACTGGCCATGTGAAAGAAAGGTAAAGAAACCACTCAGGTGAAGAACTA	8043
Oy	2895	TTGATTAAGACAAATTAGCGCGCGGTGAATTTGAAGAAAGGCTAAGATGATGACACTTGGAGT	2954
Db	8044	TTGACACACCTGTGCTATCAAAAGCTCAAAATTCACAAAGTGTCTCAGCATATGACATGAGGT	8103
Oy	2955	ACGGGACGTTCCCGACGAATGAAATTCAGACAGCTGCAGTACACGACGACAAACAC	3014
Db	8104	TGCAACAGTTTCCGGTCAACATGAAAGTAGGCGGTTCACTTAACCAAGTGAACACCTG	8163
Oy	3015	CGGGCTTTCACACTGGACCAACGCGCGCATGTCAGTATGAGATGGAGATTTACCGTAC	3074

Db 8164 AAGGGTTCTACAACCTGCGACCAAGGAGCGGTGCTGATATAGTGAAGGAGGAGATTACATCC 8223
Qy 3075 CGAGAGAGTGGGCGGGAAGAGCGA CAGCGGAAGACCGATCTCTGGAACAAGAGGAGAG 3134
Db 8224 CCGCGGAGTGAAGAGGAG 8283
Qy 3135 TTGTGGCTATTTGTTAGAGAGGTGCAAAATGAGGGGACCGGTACCGGCTTTTCAGTGTGCA 3194
Db 8284 TTGTGCGCATAGTCTCGGAGGGGCTGATGAGGGAACAAGACCGCCCTTTTCGTCGTCA 8343
Qy 3195 CTGGAACCAAGAAAGGGGTGACCATTTAGGGATACCCCGAAGTTTGTGAACCTGTGT--- 3251
Db 8344 CTTGGAATAGCAAAAGGAGAGCAATCAAGCAACCCCGAAGGAGCAGAGAGTGTGTG 8403
Qy 3252 -----CACTAGTTTACAGCGCTATGCGTCTTTCGATGTGCACTGTTCCATGCGACAAAC 3305
Db 8404 CTGCACTAGTGTACCGGCTATGCTGTGTTGAAAAGTGAAGCTTCCATGCAATCGCC 8463
Qy 3306 CACCGGTGTCTATTCACTGACGCGAAGAACGAACTGACGTCGTCGAGAGAAAGTGTG 3365
Db 8464 CGCCACATGTGTAACCCGCGAAGCATTCAGAGCTCTGACATCTCTCGAAGAAAGTGA 8523
Qy 3366 ACAATTCAAATTTAGACACGCTGCTGAGAAAGCTTGAAT--GTCCATCACCCGCGC 3422
Db 8524 ACCAGAGGCTACGACACCTGCTCAAAGCATATTTGCGGTGCGGATGTCGCGCAAGA 8583
Qy 3423 CCAAGCAAGCACTTACGATGATCTTCACTGACAGTCCCTACCTGGGGTTCGCGCT 3482
Db 8584 GTAAAGAAAGCTGCTGACGCTTACCTTACCAAGCCGTAATTTGGGACATCTCTGT 8643
Qy 3483 ATTGCAGACACTCAACGCGGTGTTTCAAGCCCAATTAATTAAGAAAGTGTGGAAGT 3542
Db 8644 ACTGTCACCTATCTGAACCGTCTTACCGCGATTAAGATGACAGAGTCTGGATGGAAG 8703
Qy 3543 CTGATGATGATGATGATTAAGATTCAGGTCTGCGCAAAATTTGGCTAACATCAGGAGCA 3602
Db 8704 CGGAGCAACACCACTAGCAGATACAGACTTCGCGCAAGTTGATGATACGCAAAAGGAG 8763
Qy 3603 CTGCGGATGTCAACAAATTCGTTACATGCTTTTGAACAGACCAATGATCAAGAGAG 3662
Db 8764 CAGCAAGCTCAAAATAGTACCGCTACATGCTGCTGAGAGAGATCAATCTGCAAAAGAG 8823
Qy 3663 ACAGATGAGAGAAATAGCTATCAGACACTGGAACCTGCGCTGTGTCGCAACAAG 3722
Db 8824 GCACATGATGATGATCAAGATCAGACCTCAGAGACCGTGTGAAGGCTTGAATCAAAAG 8883
Qy 3723 GGTACTCTCTGTTAGCTCAATGTCTCTCAAGGTGACAGTGTAAACCTGATACAGAGCG 3782
Db 8884 GATACTTTCTCTCGGAAGTGTCTCCAGGGGAGAGGCTTAAGCATAGCGAGTA 8943
Qy 3783 GAGCATCTGAGATTCATGCAACCGTGAAGAAAATCAGAGAGAAAGTTTGTGCTGAGAG 3842
Db 8944 GCAACTCAGCAACGTATGCACAATGAGCCCGCAAAATTAACAAATTTCTGAGAGCGG 9003
Qy 3843 AGGATTACTTTGTTCCACCGCTCAATGAAAGCTGTAAGTGAACGTTTACATCACT 3902
Db 9004 AAAAATATGACTTACTCTCCGTTTCAAGTGAAGAAATTTCTTTCAGATGTGACCGTCC 9063
Qy 3903 TGAAGAGACGTCTGCGGCTACATTAACCATGACAGGCAAGCCCAACGCGTATTAAGT 3962
Db 9064 TGAAGAAACAAACCGCGCTACATCACTATGCAAGGCGGAGCGGACGCTTATCAAT 9123
Qy 3963 CCTATCTGAGAGAGAGCTGAGCGGAAGTGTACATTAACCACTTCTGCAAGAAAGTCA 4022
Db 9124 CTTATCTGAGAGAAATCATCAGGAAAGTTTACGCAAGCAACCATCTCGGAGAAAGCATTA 9183
Qy 4023 CTTACGAATGTAAAGTGTGCGACTACAGCAGAGTATGCTGAGCAGGGAAGAAAGATGA 4082
Db 9184 GGTATGAGATGCAAGTGTGCGGCTTACAGACCGGAACGTTTACGACCGCTTACGAAATCA 9243
Qy 4083 ACGGCTGCACTTAAGCAAAACAGTCAATTTGCTTACAGAGCGCAACAAAGAAATGGTCT 4142
Db 9244 CGGGCTGCAACCGCATCAAGAGTGTGCTGCTTATTAAGAGCAACCAAGAAAGTGGTCT 9303

Qy 4143 TCAACTGCGCGGATCTTATTAAGGACACAGACCACTAGTGAAGTAAATTGACATTC 4202
Db 9304 TCAACTGCGCGGATCTTATTAAGGACACAGACCACTAGTGAAGTAAATTGACATTC 9363
Qy 4203 CATTCGCTTGAACCGACAGTGTGCGCGGTTCCGTTAGTCAACGCTTACAGTCAAG 4262
Db 9364 CTTTCAAGCTATCCCGAGTACCTGATAGTTCCTGTTGCCCAAGCGCGCAAGTATGAC 9423
Qy 4263 AGTGTGAAGGATCACTCTTCAACCTGATGCAATGCAAGCAACATTTGCTGACAAAG 4322
Db 9424 ACGGCTTGAACATCACTGCTTCAATTAAGACAGACCAATCTGATGCTTCAACACCA 9483
Qy 4323 GAAATTTGGGCTGAGACAGCAACGACAGAGATGATTAACAGGCTTACATCAGCA 4382
Db 9484 GGAAGCTGAGGCAACCCGGAACCAACCACTGAATGATATGGAACACGCTTAAGA 9543
Qy 4383 ATTTTCTGTGGGCGAAGAGGCTGAGTACGATAGGGGTAAACCATGAAACGATCAGAG 4442
Db 9544 ACTTCAACGCTGACCGAGATGAGCTGGAATACATATAGGGCAATCAAGAACAGTAAAGG 9603
Qy 4443 TCTGGGCGGAGAGTGTGACACAGGCGACCAATGATGAGCGGATGATCATATCC 4502
Db 9604 TCTATGCCCAAGATGTGCAACAGAGACCTTCAAGATGCGCACACGAAATAGTACAGC 9663
Qy 4503 ACTATTAATCATCGGATCAGTCACTGATGATGATGATGATGATGATGATGATGATGATG 4562
Db 9664 ATTACTATCATCGGATCAGTCACTGATGATGATGATGATGATGATGATGATGATGATG 9723
Qy 4563 TCTGTGAGGACATGATCATCAGCAGCTTGCATCGCCAAAGCAAGAGAGATGCTGTA 4622
Db 9724 TGATGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9783
Qy 4623 CGCATACGCGCTTGTGACCGAAGCAAGTAAACCAAGATGATGATGATGATGATGATGATG 4682
Db 9784 CGCATATGCGCTGCGCCCAATGCGATGATGATGATGATGATGATGATGATGATGATGATG 9843
Qy 4683 TTGGGCAACCAAGGCTGAACATTTGGAAGAACTTTGAACCACTGATGATGATGATGATG 4742
Db 9844 TTAGTGTGCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9903
Qy 4743 AACGTTTCTGTGGGACAGTGTGATCTCTGTGAGCGCTTGTATGTTCTGTTCCCT 4802
Db 9904 AGCGTCTTCTGTGGGACAGTGTGATCTCTGTGAGCGCTTGTATGTTCTGTTCCCT 9963
Qy 4803 GCTTTATGCTGATGCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 4862
Db 9964 GTTGCTCATGCTGCTGCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 10023
Qy 4863 CTTTCGAACATGCAACCTGTCGAATGTTCCGGGATGCTGCTGATGATGATGATGATGATG 4922
Db 10024 CTTTCGAACATGCAACCTGTCGAATGTTCCGGGATGCTGCTGATGATGATGATGATGATG 10083
Qy 4923 AACGCGAGTGTAGCGGCACTTAACCTGAGATCAAGTGTCTCATCGGAATTAAC 4982
Db 10084 AAAGGCAAGGATGAGCGCGCTCAATTTGAGATTAATCTGTATGTCTCTGAGAGTTTGC 10143
Qy 4983 CTTCAACTAACAGAGTGTGATCTGCAATTTCCACAGTCAATTTCTTACCAACAAG 5042
Db 10144 CTTCCACCAACCAAGAGTATTAATCTGCAATTTACCAAGTGTGTGCTTCCCTCAAAAG 10203
Qy 5043 TTAATGCTGAGGCTCTCTGAGTGAAGAGATCTTCAAAAGCGGATTAACATGCGCGC 5102
Db 10204 TCAAGTGTGAGGCTCTCTGAGTGAAGAGATCTTCAAAAGCGGATTAACATGCGCGC 10263
Qy 5103 TTTTGGGCTGTGTACCTTTATGTTGGGAGGCGCAATGCTTCTGTGACAGTGA 5162
Db 10264 TCTTTGAGGGGTGTACCTTTATGTTGGGAGGAGCAATGTTTGGCAAGTGA 10323
Qy 5163 AACACAACTGAGTGAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 5222
Db 10324 AACACCAAGATGAGTGAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 10383

QY 5223 TCGCACTAAAGTTTACACAGCTCTCTGAAAGTCGCGCTCGTATAGTATACGGCAACA 5282
 Db 10384 AGGGGATTTAAGTGTACATCTCCCGGATGAAAGTAGAGCTCGTATAGTATACGGCAACA 10443
 QY 5283 CCACCGCGCACCTGGATACGTTTGTCAAATGCGCTACCGCAGGTTCTCACGGGACCTGA 5342
 Db 10444 CTACACAGTTTCTAGATGTGTAGTGAACGGAGTCAACACGAGAACGCTTAAGAACCTGA 10503
 QY 5343 AGGTCATATGACAGGCGCCGATATACAGCCGTTTTTACACCTTTGACCATTAAGTCTGATCA 5402
 Db 10504 AAGTCATATGCTGACCAATTTTCAGATTTTACACCATTTGATCACAAGGTCTGTTATCA 10563
 QY 5403 GAAAGGGCTTGTATACACTACGACTTCTCGATATGAGTATGAAACGAGAGCGT 5462
 Db 10564 ATCCGCGCTGTGTATACACTATGACTTTCCGGAATAGCGAGCATGAAACGAGAGCGT 10623
 QY 5463 TCGGCGCATTTCAAGCATCTCTGCTTGTATGCTATACAGATATGAGCCGCGACTGACATAC 5522
 Db 10624 TTGGAGACATTCAGACTACCTCTTGAAGTCTGACTAGCAAGAACCTCATGCGCAGCAGACATTA 10683
 QY 5523 GCGTGTGAAGCCCTTCTGTGAAGAACATCCGCTCCCTACACCGCAAGCATATACGGGT 5582
 Db 10684 GCGTACTCAAGCCTTCCGCAAGAACGTCATGTCCTGATACAGAGCGCAGCATCTGAT 10743
 QY 5583 ATGAAATGTGGAAGAAACACTCAGGACGACCCCTGCAAGAAACAGACACATTTGGATGTA 5642
 Db 10744 TCGAGATGTGGAAGAAACACTCAGGCGCCGACATGCGAGAAACCGCCCTTTTGGGTGCA 10803
 QY 5643 AAATTTGAATGAGACCTCTGCGAGCGTCTAACTGTCTTACGCGCAATCCCTATCTGGA 5702
 Db 10804 AGATTTGAGTCAATCCGCTTGAGACGGGTGAGCTGCTCATACGGAACATTTCCATTCTA 10863
 QY 5703 TTGACATCCCTGATGACGCTTTTGTGATGATCATAGAAATCACAACATTTTAAAGTGA 5762
 Db 10864 TTGACATCCCGAAGCGCTCTTATACAGGACATAGAACGCACTGCTTCAACAGTCA 10923
 QY 5763 GCTGACAGATGACGAGTCTATTTATCTGAGACTTTGTGTGTTCTTAACTTAAAGT 5822
 Db 10924 AATGTGATGTGAGATGAGTGCATTTTCAAGCGACTTCGAGGAGATGGCTTCCCTGAGT 10983
 QY 5823 ACAAGCTGACAGGAGGAGCATTTGTCAAGTTCACTCCGACTCAGACGAGCTGTTTTGA 5882
 Db 10984 ATGTATCCGACCGGAGGAGCAATGCTCTGTACATTTGACATTCAGACACACACCTCC 11043
 QY 5883 AGGAAGGACCAACATGTGCTGCGGAGCAGATPAACCTATTTTGAACATCTGA 5942
 Db 11044 AAGAGTGAAGCTTCTGCTGAGAAAGGCGGTGACATCTTCACTTCAACCGCGA 11103
 QY 5943 GCCCACAAGCAAAATTTATAGTTTGTGCTATGCGGCAAGAACTGACCTGCAATCTGAT 6002
 Db 11104 GCCCACAAGGCAATCTTATTTGTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11163
 QY 6003 GTAACACCGCGCGACCAATTAATGAGAACCAATTAAGTCAACCAAGAAATTCAGG 6062
 Db 11164 GCAAAACACACGATCATATCTGTAGACACCCGCGCAAAATAAGCAAAATTTCCAAG 11223
 QY 6063 CGGAGATTTCCAAACATTTTGAACCTGCTGTGACATCTTTGGGGAGACATATCCC 6122
 Db 11224 CGGAGATTTCCAAACATTTTGAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11283
 QY 6123 TCATTTGTGTGAGCTTATAGT 6182
 Db 11284 TATTAATTTATGAGCTTATATGATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11343
 QY 6183 GACTGAGCGCGGACACTGA-CATAGCGGTAAACCTCATGTTCTTCCGAGGAGCGTGTGT 6241
 Db 11344 GACCGCTAAGCGCCCAATGACCCGACACGACAAACCTCATGTTCTTCCGAGGAGCTGATGT 11403
 QY 6242 GCATTAATGCGCGCGCGCT-----TGACACT 6288
 Db 11404 GCATTAATGCGCGCGCGCTGT 11463
 QY 6289 AAAAAGCTGATGTATTTCCGAGGAAGCACAGTCAATATGCTGTGCAAGTGT 6318

Db 11464 AAAAAGCTGATGTATTTCCGAGGAAGCGCAGTCAATATGCTGTGCAAGTGT 11513
 RESULT 3
 AAV3473
 ID AAV3473 standard; cDNA; 11663 BP.
 AC AAV3473;
 XX
 DT 10-DEC-1998 (first entry)
 XX
 DE South African Arbovirus strain No. 86 cDNA sequence.
 XX
 KW Bone marrow cell; alphavirus; South African Arbovirus strain No. 86;
 KW S.A.ARB6; reverse transcriptase-PCR; nsP1; nsP2; nsP3; nsP4; capsid;
 KW E3; E2; 6K; E1; growth hormone; growth factor; interleukin; cytokine;
 KW chemokine; enzyme; ribozyme; antisense oligonucleotide; ss.
 XX
 OS South african arbovirus
 XX
 FH Location/Qualifiers
 FT 1..59
 FT 5'UTR
 FT /+tag= a
 FT CDS 60..7662
 FT /+tag= b
 FT /product= "Non-structural polypeptide"
 FT 7608..11345
 FT /+tag= c
 FT /product= "Structural polypeptide"
 FT 11346..11663
 FT 3'UTR
 FT /+tag= d
 FT
 XX W09836779-A2.
 XX 27-AUG-1998.
 XX
 XX 18-FEB-1998; 98WC-US02945.
 XX 19-FEB-1997; 97US-0801263.
 XX (VUNC-) UNIV NORTH CAROLINA.
 PA
 PI Davis NL, Johnston RE, Simpson DA.
 XX WPI; 1998-495361/42.
 DR P-PSDB; AAW70460, AAW70461, AAW70466, AAW70467, AAW70468, AAW70469,
 DR AAW70470, AAW70471, AAW70472.
 XX
 PT Expression of heterologous RNA in bone marrow cells - using a
 PT recombinant alphavirus comprising a promoter operable in bone marrow
 PT cells operably associated with heterologous RNA.
 PS
 PS Example 2; Fig 1A-1C; 68pp; English.
 CC The invention provides a method for introducing and expressing
 CC heterologous RNA in bone marrow cells using alphavirus vectors.
 CC The South African Arbovirus strain No. 86 (S.A.ARB6) is an
 CC alphavirus whose cDNA sequence was determined from uncioned reverse
 CC transcriptase-PCR reaction fragments amplified from the virion RNA.
 CC The S.A.ARB6 cDNA encodes nonstructural and structural polypeptides.
 CC The nonstructural polypeptide is, presumably, post-translationally
 CC modified into four different protein products, namely nsP1 (AAW70460),
 CC nsP2 (AAW70461), nsP3 (AAW70466), nsP4 (AAW70467) proteins. The
 CC structural polypeptide is, presumably, post-translationally modified into
 CC five different protein products, namely capsid (AAW70468), E3
 CC (AAW70469), E2 (AAW70470), 6K (AAW70471), E1 (AAW70472) proteins. The
 CC S.A.ARB6 cDNA was used in the method of the invention. The inventors
 CC claim the transformed bone marrow cells are useful for expressing a
 CC protein or peptide suitable for protecting the subject against a disease
 CC such as a microbial, bacterial, protozoal, parasitic or viral disease.
 CC The transformed bone marrow cells are also claimed to be useful for
 CC expressing proteins and peptides such as hormones, growth hormones,

CC growth factors, interleukins, cytokines, chemokines, enzymes, ribozymes
or antisense oligonucleotides.

XX Sequence 11663 BP; 3291 A; 3105 C; 2907 G; 2360 T; 0 other;

Query Match 36.9%; Score 2396; DB 19; Length 11663;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3751; Conservative 0; Mismatches 1985; Indels 74; Gaps 9;

QY 580 ACTGACGATGTAAGGAGGCTATATTTCTCTCGGAAACAGGCCAAGTCACTTC 639
DB 5707 ACTGCTAACCGGGGTAGGTGGTACATATTTTCACGACACAGGCCCTGGGCACTTGC 5766
QY 640 AACGAAATTCAGTACGTCAATGTAACTACAGAACTATATTTGGATCGGGCCGTCAATG 699
DB 5767 AAAAGAAATCGCTTGTGCAAGACCAAGCTTACAGAACCACTTGGAGGCAATGTCTCG 5826
QY 700 AGAAGTATTAAGCCCGCCCTCGATCTCGAAAGAGAAATGTTACAGAAAGAACTGC 759
DB 5827 AAAGATCTACGCCCCGGGTCTGCACAGTGGAAAGAGAACAGCTCAAACTCAGGTACC 5886
QY 760 AATTATGGCTCTTAAGGAAATTAAGACAGTATCATACGAAAGTAAAGAAATAGA 819
DB 5887 AGATGATCCCAACCGAGCCAAACAAAGCAGGTACAGTCTCGAAAGTAGAAACAGA 5946
QY 820 AAGCAATTACAGCGGAGCACTCATTTCTGATGGGACATATCTATCATCAGAAAGTGA 879
DB 5947 AAGCATTAACCACTAGAGCATCTTTTCAAGGCTACGACTGTATTAATCTGCGACAG--- 6003
QY 880 ATCTCTGAGGTGTACAGAGTCAATATCTGTACCAATCTACTCTGCAACGATATTA 939
DB 6004 ATACGCCAGAAATGCTATAGATCACCTACCGAAACACATCTATTCAGAGGTATACAG 6063
QY 940 ACAAGTTACATCTGCAAGAGTGGGGTTAAACCTGCACTTATTCAGAGAAATT 999
DB 6064 CGAATCTCTGACCCCAAGTTGCTGTAGCTGTTTGAACAATATCTGATGAGAAATT 6123
QY 1000 ACCCTACAGTACCGAGTATTTGTATACAGATGAATACGATCGTATCTTGCATGATGG 1059
DB 6124 ACCGACGCTGACATCTTATCAGATCACCAGAGATGATGATCTTATCTTGGATATGGT 6183
QY 1060 ACGGCGCATCTGCTGTCTAGTATACAGCACTTTTGTCCGGCTAACTGAGAACTAC 1119
DB 6184 ACGGACAGTGGCTTGCCTAGTATCTGCAACTTTTGGCCCGCAAGCTTGAATTTACC 6243
QY 1120 CAAGAGACATAGCTATTTTGCAGCGAGATTAAGATCAGCCGTCCCATGCTTATACGA 1179
DB 6244 CGAAAGACACGAGATATAGAGCCCCCAATCCGAGTGGGTTCCATCAGAGATGACAGA 6303
QY 1180 ATACATTAACAATAATGTATTTGGCTGACGATTAAGAAAGAAATTGCAACGTTACCCAAATGC 1239
DB 6304 ACACCTTTCACAAACGTGCTCATTTGCCGACATTAAGAAAGAACTGCAACGTCACACAAATGC 6363
QY 1240 GAGATTTACTGCTTCTTATGATTCGGGCAATTAATGTTGATTTTCAAGAAATACGAT 1299
DB 6364 GTGAACCTCCAAACATGAGTACTAGGACATTTCAAGTTGAATGCTTTGAAAAATATGAT 6423
QY 1300 GCATGATGATGATCTGGGATACCTTTCCGGATTAACCTTATTCGGCTTAATCAAGAAAG 1359
DB 6424 GCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6483
QY 1360 TTACGATATATGTCAGAAAGCTGAAGGCGGAAAGCAGACAGATTTGTTGGAAATATC 1419
DB 6484 TTACGATATATGTCAGAAAGCTGAAGGCGGAAAGCAGACAGATTTGTTGGAAATATC 6543
QY 1420 ATATATTAACCCGTTGACAGAGATACCAATGATCAATTCGTCAATGATCTTAAGAGAG 1479
DB 6544 ATATATTTGCTCCATTTGCAAGAAAGTCCATGATGATGATGATGATGATGATGATGAT 6603
QY 1480 ATGTAAAGTTACTCCGCGACGAAACATACAGAGAGCGGCTTAAGTGCAGGTTATTC 1539
DB 6604 ACGTAAAGTTACTCCGCGACGAAACATACAGAGAGCGGCTTAAGTGCAGGTTATTC 6663

QY 1540 AGGCTGACATCCCTTGTCTACCGCTTACCTTTGCGGAGTCCATCGGAAATTAATGCTCA 1599
DB 6664 AAGCGGAGAAACCTCTGGCGACCGCTTATCTATGCGGATCCACCGGAGTTAGTGGCA 6723
QY 1600 GACTGATGCGGTGCTTGTGCAAAATATCCATCTCTTGTGACATGTTCAGCGGAATTT 1659
DB 6724 GCTTACAGCGGTTTGTGTCACCAATTCACAGCTCTTGTGAATGTCCGCGGAGACT 6783
QY 1660 TTGAATGATATTTGCTGAATTTCCACGCGGAGACCGCATTTGGAAACGACATCG 1719
DB 6784 TTGAATGATATATGACAAACATTCACAGAGGTGACCGGATCTGGAAGACGATATCG 6843
QY 1720 CGTCTTTGATTAAGGCAAGACAGCGTATCGCATTTCCGCGGTGATGATTCCTTGAG 1779
DB 6844 CCTGCTTGCAGAAAGCCAAACAGCATATGCGGTTAACCGGCTGATGATCTTGGAG 6903
QY 1780 ACTTAGGTGACCAACCGCTCTTAATTTGATGAGGCGGCTTCCGCAATATCATAT 1839
DB 6904 ACTGCGGTGAGCAACACATCTGATCTGATGAGTGGCGCTTTGGAAATATCATAT 6963
QY 1840 CTGTGACCTTACCAAGGAAACGAGTTAAATTTGGTCCATGATGAATCCGATATCT 1899
DB 6964 CCACCATCTGCGCAGCGGTACCGGTTCAATTTGGGCGATGATGAATTCGGAATGT 7023
QY 1900 TCTTAACGCTGTTGTCAACACATGATATATATGATGATGATGATGATGATGATGATGAT 1959
DB 7024 TCTTACGCTCTTTGTCAACACATGATGATGATGATGATGATGATGATGATGATGATGAT 7083
QY 1960 AACGGTTAACAGGTCAGCGGTGCGCGCTCTATTCGCGAGATTAACATATGATGATG 2019
DB 7084 AGCGGCTTAAACCTCCAAATGTCAGATTTATTCGGGAGACAACTTATACAGGAG 7143
QY 2020 TCGTCTCCGACCTTGAATGCGGAGAGATCCGCACTTGGCTGAACATGGAATTAATA 2079
DB 7144 TAGTATCTGACAAAGAAATGCTGAGAGTGTGCACCTGCTCAACATGAGGTTAAGA 7203
QY 2080 TTATGATGACGTTATTTGATTTCAAGACCCCTTCTGTGGGGGATTTATCTCGTGG 2139
DB 7204 TCATTTAGCGAGTATCGGCGAGAGACCACTTATCTTGGGATGATTCATCTTGCAG 7263
QY 2140 ACCAGATTAACGAGCAGACCTGCAAGTGGAGACCTCTAATAAGGCTTTTAAAGCTTG 2199
DB 7264 ATTCGTTACTCCACAGCGGTGTGCGTGGGAGACCTTTAATAAGGCTTTTAAAGTGG 7323
QY 2200 GAAACCATTTCCGATGATATACCAAGCTGCGACCGCGCGGCACTGATGATG 2259
DB 7324 GTAAACCGCTCCAGCGGACGATGAGCAAGACAGAAAGCGCCTCTGCTAGATG 7383
QY 2260 AAGCAATGCGATGGAACGAATTTGGAATTACGAGCGATTTGTAAGGCGGTAGAATCCA 2319
DB 7384 AAACAAAGCGGTGTTTGAATGATGATTAACAGACCTTTGACAGTGGCCGTGCACTC 7443
QY 2320 GATACGATATATCTGCGAGCGCTGATCATCAGCTCTGTGCAAGTTAACCGAAAGCG 2379
DB 7444 GGTATGAGTATGACAAATCATCACCTGTCTGCTGGCATTTAGAACTTTTCCAGAGCA 7503
QY 2380 TTAAAGACTTCAAGACATTAAGAGGAGCCCAATCACCTTTAGCGCTGATTAATG 2439
DB 7504 AAAGAGCAATTTCAAGCATCAGAGGGAATTAAGCAATCTCTACGCTGTCTTAATAGT 7563
QY 2440 TGACGTAGTAGA-----CAGGACCTTACCGAGCGGAGATGTTTCCA 2482
DB 7564 CAGATATGATATTTCACTGATTAATACCAACACACACCATGATTAATGAGATTTCT 7623
QY 2483 TACCTCAGCTGAACCTTTCACACAGTTAACCTTAACATTCGATGCTTACCGAGATCCA 2542
DB 7624 TTAAATGCTGCGGCGCGCGCTTCCAGCGCCCACTGATGAGGCGCGGAGAA 7683
QY 2543 AACCTCTTAGCGCGCTGAGGCGCTTGGCGCCCGCTGCTGCTCAATGAAGAT 2602
DB 7684 GAGGACAGCGCGCCGATGCTGCGCAATGAGCTGAGCTTCCCAATTCAGCACTGA 7743
QY 2603 CTTAGAGGTGATGATCAATCTTCAAAAC--GATCACTTAATCCGCGCGAG 2660

Dh 7744 CCAAGCCGTCAGTGCCTTACTGATTTGGACAGGCACTGAGCCCTCAAAACCCACAGCCAC 7803
Qy 2661 GTCCACCGCCAAAGAGAGAGAGTGTCTTAAAGCCAAAACCTTACTGACCTTAAAAA 2720
Db 7804 GCCCGCCGCCCGCAGAGAGAGAGAGGCGCCAAAGCAACCCAGAGCCGAAAGAAACCA 7863
Qy 2721 AGAGAGAGCAAGCAAGAGAGAGCAAAAC--GCAAGCCCTAAACCAAGGGAACAGACAGT 2777
Db 7864 AAGACAG 7923
Qy 2778 TGTGTATGAAGTTGAGTGGAGTCAGACAGATTTCCGATC--ATGCTGAACGCCCAAGTGA 2834
Db 7924 TGGACATTAAGTTGAGAGCCGACAGACTGTTGACAGTCAAAATAGAGAGAGAGATGCA 7983
Qy 2835 ATGATATGCTGCTGCTGTTGCGAGAGAGGCTGATGAACCACTCCAGCTTGAAGAAAA 2894
Db 7984 TCGGAGCAGCACTGGCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8043
Qy 2895 TTGATTAATGAGCAATTAAGCGCGCGTGAATTAAGAGAGGCTAGAGATGATGAGCTTGA 2954
Db 8044 TTGACACCTGTGCTATCAAAAGCTCAAAATTCACCAAGTGTGACAGATGACAGATGAG 8103
Qy 2955 ACGGCGACGTTCCCGAGACATGAATTCAGACAGCTGCACTAGACAGAGCAAAACAC 3014
Db 8104 TCGACAGTTGCCGCTCAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8163
Qy 3015 CGGCTTCTACACTGTCACCAAGCGCGAGTTCAGATGAGAGATGAGAGATTTACCGTAC 3074
Db 8164 AAGGTTCTACAACTGGACACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8223
Qy 3075 CGAGAGAGAGTGGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3134
Db 8224 CCGCGAGATGAG 8283
Qy 3135 TTGTGCTATTTGTTCTAGAGAGTGCATTAAGAGAGAGAGAGAGAGAGAGAGAGAG 3194
Db 8284 TTGTGCGATAGTCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8343
Qy 3195 CTTGGAACAG 3251
Db 8344 CTTGGAATGAG 8403
Qy 3252 -----CACTAGTTAAGCGGCTATGCGGCTTTGGAATGTCAGCTTCCATGCGCA 3305
Db 8404 CTGACACCACTGTCAGCGGCAATGCTGCTTTGGAAGAGTGAAGCTTCCATGCAATG 8463
Qy 3306 CACCGTGTCTATTTCACTGACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3365
Db 8464 CGCCACATGCTACACCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8523
Qy 3366 ACAATCCAAATTAAGACAGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3422
Db 8524 ACCACAGAGGCTTACGACACCTGCTCAAGCCATATTCGCTGAGAGAGAGAGAG 8583
Qy 3423 CCAAGAGAGATTAACGATGATTCACAGTCAAGAGAGAGAGAGAGAGAGAGAGAG 3482
Db 8584 GTAAAAAGAGGCTACGACAGCTTACCTTGACAGCGCCGATCTTGAGGACAGAGCT 8643
Qy 3483 ATTGACAGACTCAACGCGTGTTCAGCGCCCAATTAATTAAGAGAGAGAGAGAGAG 3542
Db 8644 ACTGTACCACTATGTAACCGTGTTCAGCGCCGATTAAGATGAGAGAGAGAGAGAG 8703
Qy 3543 CTGATGATGATTCGATTAAGATTCAGAGTCTCGGCAAAATTCGCTTACATGAGAG 3602
Db 8704 CGGAG 8763
Qy 3603 CTGCGAGATGACCAAAATTCGCTTACAGTCTTTGACAGAGAGAGAGAGAGAG 3662
Db 8764 CAGCAAGCTCAAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8823
Qy 3663 AACAGTATGAG 3722
Db 3722

Dh 8824 GCACATGATGACATCAAGATGACAGCACTCAGAGCCGTGTAGAGAGAGAGAG 8883
Qy 3723 GGTACTTCTGTATAGTCAATGTCCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3782
Db 8884 GATACCTTCTCCCGAG 8943
Qy 3783 GAGCATCTGAGAGATTCATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3842
Db 8944 GCAACTGAG 9003
Qy 3843 AGAGAGATCTTGTTCACACCGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAG 3902
Db 9004 AAAAATATGACATACCTCCGTTACAGGATGAAGATTCCTTCCAGAGAGAGAGAG 9063
Qy 3903 TGAAGAGAGAGTCTGCGGCTATCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3962
Db 9064 TGAAG 9123
Qy 3963 CTTATCTGAG 4022
Db 9124 CTTATCTGAG 9183
Qy 4023 CTTATGAG 4082
Db 9184 GTAG 9243
Qy 4083 ACGGCTGACATTAAG 4142
Db 9244 CCGGCTGACAG 9303
Qy 4143 TCAATCGCGAG 4202
Db 9304 TCAATCGCGAG 9363
Qy 4203 CATTCGCTTGAACAG 4262
Db 9364 CTTTCAAGCTGATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9423
Qy 4263 AGTGTTCAG 4322
Db 9424 ACGGCTTAAACATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9483
Qy 4323 GAAATTTGGGCTGCGAG 4382
Db 9484 GAGAGCTAGGAG 9543
Qy 4383 ATTTTCTGTGGGCGAG 4442
Db 9544 ACTTCAACGTCGACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9603
Qy 4443 TTTGGGCGAG 4502
Db 9604 TCTATGCCAAGAGTTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9663
Qy 4503 ACTATTCATTCGAG 4562
Db 9664 ATTATTCATTCGAG 9723
Qy 4563 TCTGTAG 4622
Db 9724 TGATGATTTGCGTAACTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9783
Qy 4623 CGGCTTACGCGCTTTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4682
Db 9784 CGGCTTACGCGCTTTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9843
Qy 4683 TTGCGCAACCAAG 4742
Db 9844 TTGAGTGGCTATGAG 9903
Qy 4743 AACCGTTCTCTGGGAG 4802
Db 9904 AGCGTTCTTCTGGGAG 9963

QY 4803 GCTTTTCATGCTGATGCTTTTATTTATGTTGACAGCGCTGCTGGGGAAGTAAAGC 4862
 DB 9964 GTTGCTATGCTGCTGCTTTTATTTAGTGTTCGGGCGCCCTACCTGGGGAAGTAAAGC 10023
 QY 4863 CCTTGAACATGCGAACCACTGTGCCAATGTTCCGGGATCCGTAATAAGCGTTGTG 4922
 DB 10024 CCTAGGAACATGCGAACCACTGTGCCAATGTTCCAGATATCCGTAATAAGCACTTTGTG 10083
 QY 4923 AACGGCAGGTAAAGCGGCACTTAACCTGAGATACAGGTGCTCATTCGGAATTAACAC 4982
 DB 10084 AAAGGCGAGGTAAAGCGGCACTTAACCTGAGATATCCGTAATAAGCACTTTGTG 10143
 QY 4983 CTTCAACTAACAGAGATGATGATCTGCAATTCACACAGTCACTTCTTACCAACA 5042
 DB 10144 CTTCAACCAACAGAGATGATGATCTGCAATTCACACAGTCTGCTTCTTCTTAAAG 10203
 QY 5043 TTAATGCTGCGGGTCCCTCGAGTGCAGAGGATCTCAAAAGCGGATTAACATGCCG 5102
 DB 10204 TCAGATGCTGCGGGTCCCTCGAGATGTCAGCCGCGCTCACAGCACTATACCTCAAG 10263
 QY 5103 TTTTGGGGGTGTATACCTTTCATGTGGGAGGCGCACATGCTTCTGTGACATGAGA 5162
 DB 10264 TCTTGGAGGGGTGTATACCTTTCATGTGGGAGGCGCACATGCTTCTGTGACATGAGA 10323
 QY 5163 ACACACAATGATGAGGAGTACGTCGAGTTCGTCAGATGCACTATATGATCAAGG 5222
 DB 10324 ACAGCAGATGATGAGGAGTACGTCGAGTTCGTCAGATGATGTCGCGATGACACAGG 10383
 QY 5223 TCGCATTAATAAGTTCACACAGCTGCTGTAAAGTCCGCTGCGATATGATATACGCA 5282
 DB 10384 AGGCGATTAAGTTCACACAGCTGCTGTAAAGTCCGCTGCGATATGATATACGCA 10443
 QY 5283 CCACCGCGCACTGATGATGCTTTCATGAGCGTCAAGCGAGTCTTCAAGGACCTGA 5342
 DB 10444 CTACCGATTTCTTATGATGATGATGAGGAGGAGTCAACAGGAACTGTAAAGCTGA 10503
 QY 5343 AGGTATAGCAGAGGCGGATATGAGCGGCTTTTTCACCTTTCAGCATAGAGTCTGATCA 5402
 DB 10504 AAGTATAGTGTGAGCAATTTTACGATGTTTACACCATTCGATACAGAGTCTGATCA 10563
 QY 5403 GAAGGGGCTTGTATTAACAATGCACTTCCCTGATATGAGCTATGAACAGGAGCGT 5462
 DB 10564 ATCCGGGCTGTGTATTAACAATGCACTTCCCTGATATGAGCTATGAACAGGAGCGT 10623
 QY 5463 TCGGGAATTTACAGATCTTCTGCTGATGCTACAGACATATGACCGGCACTGATAC 5522
 DB 10624 TTGGAGATTTACAGATCTTCTGCTGATGCTACAGACATATGACCGGCACTGATAC 10683
 QY 5523 GGCTGCTGAGGCTTCTGTCAAGAACATCAAGTCCCTCAAGCCCAAGGATATGAGGT 5582
 DB 10684 GGCTATCAAGGCTTCTGTCAAGAACATCAAGTCCCTCAAGCCCAAGGATATGAGGT 10743
 QY 5583 ATGAATGTGGAAGAACAATCTAGAGACGACCCCTCAAGAAACAGACCATTTGATGTA 5642
 DB 10744 TCGAGATGTGGAAGAACAATCTAGAGACGACCCCTCAAGAAACAGACCATTTGATGTA 10803
 QY 5643 AAATTTGAAGTGAAGCTTCTGCGAGGCTTCACTGCTTACGCGCAATCCCTATCTGA 5702
 DB 10804 AGATTTGAGTGAAGCTTCTGCGAGGCTTCACTGCTTACGCGCAATCCCTATCTGA 10863
 QY 5703 TTGACATCCCTGATGAGCTTTTGTGAGATCATCGAATCCCAATTTTGAAGTGA 5762
 DB 10864 TTGACATCCCAAGCTTCTTATCAAGACATCAAGATGACACATGCTGCTCAACAGTCA 10923
 QY 5763 GCTGACAGTACAGACTGATTTATTTCTGACAGCTTTGTGTCTTAACTATACAGT 5822
 DB 10924 AATGTAGTGTAGTGTAGTGAATTTATTTACAGGAGCTTTGAGAGGATGCTACCTGAGT 10983
 QY 5823 ACAAGCTGACAGGAGGAGCAATTTGCTCAAGTTCATCCCACTTCAAGAGCTGTTTGA 5882
 DB 10984 ATGTATCCAGCGCGAAGGACATGCTCTGTATCATTTGATGAGACAGCAACCTTCC 11043

QY 5883 AGAAGCAGACACATGATGATGCTGCGTAGGAGAGATTAACATTAATTTAGCATGCA 5942
 DB 11044 AAGAGTGCACAGTTTATGCTTCTGAGAAAGAGAGCGGTGACAGTACACTTACGACCGCA 11103
 QY 5943 GCCCAGACAAATTTATGATTTGCTATGCGGCAAGAGTCCACTGCAATGCTGAAT 6002
 DB 11104 GCCCAGAGCAAACTTATGATGCTGTGTGTGTAAGACAAATGCAATGCAAT 11163
 QY 6003 GTAAACCAACCGGCGACACATTAATTTGAGAAACCAATTAAGTGCACCAAGATTCACG 6062
 DB 11164 GCAACCAACAGCTGATATATCGAGACACCCGCACAAAATGACCAAGATTTCAAG 11223
 QY 6063 CGGCACTTCCAAACATCTTGAAGTGGCTGCTTGAAGTCTGTTGGGAGCATATCC 6122
 DB 11224 CCGCATCTCAAAACTTCAATGAGTGGCTGTTGGCTTTTCCGCGCGCTGTGCG 11283
 QY 6123 TCATTTGTGAGACTTAATGATGTTGTGTCGAGCTATGCTTATTAACACACAGT 6182
 DB 11284 TATTATATATGAGCTTATGATTTTTCCTGACAGATGATGCTGACACAGAGAT 11343
 QY 6183 GACTGAGGCGGAGACATGCA-CATAGCGGTAAACCTGATATCTTCCAGAGAGCTG 6241
 DB 11344 GACCGTACGCCCCAATGACCCGACAGCAAAACTGATGATCTTCCAGAGAGCTGAT 11403
 QY 6242 GCATATGCGCAGCGCGCT-----TGACACT 6268
 DB 11404 GCATATATCATGAGCTGTATATTAATCCCGCTTACCGCGGCAATATAGCAACAC 11463
 QY 6269 AAAATGATATTTCCGAGAAAGACAGTGCATATGCTGTGAGTGT 6318
 DB 11464 AAAACTGACGATATTTCCGAGAAAGCGCAGTGCATATGCTGTGAGTGT 11513

RESULT 4
 AAV33476
 ID AAV33476 standard; cDNA: 11703 BP.
 XX
 AC AAV33476;
 XX
 DT 10-DEC-1998 (first entry)
 XX
 DE Sindbis virus clone TR339 cDNA sequence.
 XX
 KW Bone marrow cell; alphavirus; Sindbis virus; cytokine; capsid;
 KW reverse transcriptase-PER; nsP1; nsP2; nsP3; nsP4; E3; E2; E1;
 KW growth hormone; growth factor; interleukin; chemokine; enzyme;
 KW ribozyme; antisense oligonucleotide; ss.
 OS Sindbis virus.
 XX
 FH Key Location/Qualifiers
 FT 1..59
 FT 5'UTR
 FT
 CDS
 FT 60..7601
 FT /*tag= a
 FT /product= b
 FT /note= "Non-structural polypeptide"
 FT /note= "Internal stop codon present at nucleotides
 FT 5748-5750"
 FT 7647..11384
 FT /*tag= c
 FT /product= "Structural polypeptide"
 FT 11382..11703
 FT /*tag= d
 FT 3'UTR
 FT
 PN MO9836779-A2.
 XX
 PD 27-AUG-1998.
 XX
 PF 18-FEB-1998; 98MO-US02945.
 XX
 PR 19-FEB-1997; 97US-0801263.
 XX
 PA (UNNC-) UNIV NORTH CAROLINA.

XX Davis NL, Johnston RE, Simpson DA;
PI WPI: 1998-495361/42.
XX P-PSD; AAW70464, AAW70465, AAW70480, AAW70481, AAW70482, AAW70483,
DR AAW70484, AAW70485, AAW70486.
XX Expression of heterologous RNA in bone marrow cells - using a
PT recombinant alphavirus comprising a promoter operable in bone marrow
PT cells operably associated with heterologous RNA.
XX Example 6; Fig 6A-6B; 68bp; English.
XX The invention provides a method for introducing and expressing
CC heterologous RNA in bone marrow cells using alphavirus vectors.
CC The present sequence represents the Sindbis virus clone TR319 cDNA
CC sequence. The Sindbis virus cDNA encodes nonstructural and structural
CC polyproteins. The nonstructural polyprotein is, presumably,
CC post-translationally modified into four different protein products,
CC namely nsP1 (AAW70464), nsP2 (AAW70465), nsP3 (AAW70480), nsP4
CC (AAW70481) proteins. The structural polyprotein is, presumably, post-
CC translationally modified into five different protein products, namely
CC capsid (AAW70482), E3 (AAW70483), E2 (AAW70484), 6K (AAW70485), E1
CC (AAW70486) proteins. The Sindbis virus cDNA was used in the method of
CC the invention. The inventors claim the transformed bone marrow cells
CC are useful for expressing a protein or peptide suitable for protecting
CC the subject against a disease such as a microbial, bacterial, protozoal,
CC parasitic or viral disease. The transformed bone marrow cells are also
CC claimed to be useful for expressing proteins and peptides such as
CC hormones, growth hormones, growth factors, interleukins, cytokines,
CC chemokines, enzymes, ribozymes or antisense oligonucleotides.
XX
SQ Sequence 11703 BP; 3306 A; 3051 C; 2913 G; 2433 T; 0 other;
Query Match 36.8%; Score 2388.4; DB 19; Length 11703;
Best Local Similarity 64.7%; Pred. No. 0;
Matches 3698; Conservative 0; Mismatches 1981; Indels 41; Gaps 8;
QY 580 ACTGACGGTATGAAACCGGACCGTATTTCTCATCGGAAACAGGCCAAGTCACTTC 639
DB 5746 ACTACTCAACCGGGGTAGTGTGTACATATTTTCAGACGACACAGGCCCTG3GACCTTGC 5805
QY 640 AACGAAATCAGTACGTCATATGTAACAAAGCACTATATTTGGATCGGCGCTCCATG 699
DB 5806 AAAAGAGTCCGCTTCTGAGAACCAAGCTTACAGAACCACTTGGAGCGCAATGCTCG 5865
QY 700 AGAAGTATTCGCCCCCGCGCTCGATCTCGAAGAGAGAAATGTTACAGAAAGAACTGC 759
DB 5866 AAAAGATTCATGCCCCGCGTCTGACACGTGAAAGAGGAACTCAAACTCAGTACC 5925
QY 760 AATTATCGCCTCGAAGGAATAGAGCAGTATCATATCGAAAGTAAATATGA 819
DB 5926 AGATGATGCCACCGAGCCCAAAAGTAGTACAGTCTCGAAAGTAAATAATCGA 5985
QY 820 AAGCAATTCACGCGAGGAGCTCATTTCTGATTTGGGACATATCTATCATCAGAAGTGA 879
DB 5986 AAGCATTAACCACTGAGGAGTACTGTCAGACTACAGACTATTAATCTTCCACAG--- 6042
QY 880 ATCTGTCGAGTGTAAAGATCAATTAATCTGTATCAACATCTACTCGTCAACGGTAAATTA 939
DB 6043 ATCAGCCAGATGTCTATTAAGATCACTATCCGAAACATGTACTCCAGTACGTAACGG 6102
QY 940 ACAGTTTACATCTGACAGAGTCCGGTAAACAGTGAATTAAGTTATCCAGAGAAATT 999
DB 6103 CGAATCTCTCGATTCACAGTTCGCTGTGCTGTATCAACATCTGTGATAGAACT 6162
QY 1000 ACCCTACGATGAGCAGTATATGTATTAACAGATGATAGATCGGATTTGACATGGTGG 1059
DB 6163 ATCCGACAGTACATCTTATCAGATTAAGTACAGATGATGATCTTATCTGATATGGTAG 6222
QY 1060 ACGGCGCATCGTGTCTGTAGATACAGCACTTTTGTCCGGCTAAACTGAGAAAGTACC 1119
DB 6223 ACGGACAGTGTGCTGTGATTAAGTCACTTCTGCGCCGCTAAGTGAAGTTAAC 6282

QY 1120 CAAAGAGCATAGTATTTGACAGCAGAGATAAGATCAGCGTCCCTCATACAGA 1179
DB 6283 CGAAAAACATGAGTATAGAGCCCGCAATATCCGAGTGGCTTCCATCAGCATGACAGA 6342
QY 1180 ATACATTACAAAATGTAATTTGGCTGACAGTCTAATAAGAAATTCGAAGTCCCAATGTC 1239
DB 6343 ACACGCTACAAAATGTCCTCATTTGCGGCACTAATAAGAAATTCGAAGTCCAGCAGATGC 6402
QY 1240 GAGATTTACCTGCTTATGATTTGCGCGCATTTTAATGTTATTTTCAAGAAATACCAT 1299
DB 6403 GTGAACGCGCAACATGAGTCTCAGCCGACATTCATATGCGAATCTTTCGAAATATGCAT 6462
QY 1300 GCATGATGAGTCTGGGATACCTTTGCGATTAACCTTATGCGCTACTACAGAGAACG 1359
DB 6463 GTATATGCGATTTTGGAGAGAGTTCGCTCGAAGCCATTTAGATTAACACTGAGTTTG 6522
QY 1360 TTACGCAATATGTGACAAAGCTGAAAGGCGGAAAGCAGAGCATTTGTCGAATATCTC 1419
DB 6523 TCACCGCATATGTAGTACGTAAGTGAAGGCGCGGACATTTTGCAGAAAGCGT 6582
QY 1420 ATATATCTAAACCGTTGACAGAGATACCAATGATCAATTCGTCTATGATCTTAAGAGAG 1479
DB 6583 ATATATTTGTCCCATTTGCAAGAGTGTCTATGATATTCGTCTATGATGATGAAGAG 6642
QY 1480 ATGTCAAGTTATCTCCCGGACGCAACATATCAGAGAGCGGCTTAAGTGCAGGTTATTC 1539
DB 6643 ACCTGAAAGTTACACAGGACGCAACACAGAAAGAACCCGAAAGTACAGATGATAC 6702
QY 1540 AGGCTGAGATCCCTTGTCTACCGCTTACCTTTGCGGATTCATCGGAAATTAATGCCGTA 1599
DB 6703 AAGCGCGAAGACCCCTGCGACATGCTTACTATGCGGATTCACCGGAAATTAATGCCGTA 6762
QY 1600 GACTGAATGCGGTCTTCTGCAATATTCATATCTCTTTCGACATGTCAGCGGAAGATT 1659
DB 6763 GCGTTACGCGGCTTCTGCTTTCGCAACATTCACAGCTTTTGAATGATTCGCGGAGATT 6822
QY 1660 TTGATGATTAATTTGCTGAACATTTCCACACGCGGACCCAGTATTTGAAAGCAGACATCG 1719
DB 6823 TTGATGCAATCATATGCAAGAACCTTCAAGCAAGGCGACCCGTAATGTAAGACGATATCG 6882
QY 1720 CGTGTGTTGTAAGGGAAGCAGACGCTATCGCATTTTGGGCGTTGATGATCTTTGAGG 1779
DB 6883 CATCATTTGCAAAAGCCAGACGACGATGCGTTAACCGGTGATGATCTTTGAGG 6942
QY 1780 ACTTAGGTGACCAACCGCTCTTGAATTTGATAGAGGCGGCTTGGCAATATCAT 1839
DB 6943 ACCTGGGTGTGATCAACCACTACTGATCGATGATGAGTGGCTTTGGAGAAATATCAT 7002
QY 1840 CTGTGACCTACCTACAGAAAGAGGTTAAATTTGGTGCATGATGAATCCGATATGT 1899
DB 7003 CCACCATCTACTACGAGGATCTCGTTTAAATTTGGGCGATGATGAATCCGAAATGT 7062
QY 1900 TCTTAACGCTGTTTGTCAACACACTAGTCAATATCATGATTTGCTAGCAGATCTACGTTG 1959
DB 7063 TCTCAACACTTTTGTCAACACAGTTTGAATGCGTTATTCGACAGAGTACTAGAG 7122
QY 1960 AAGGTTAACCAAGTCAAGCTGCGGCGCTCTATCGGACGATTAATAGTGCATGTTG 2019
DB 7123 AGCGGCTTAAAGCGTCAAGATGTGCAAGCTTATGCGCAGCAACATATATCATAGAG 7182
QY 2020 TCGTCTCCGACCTTGAATGGGCGAGAGATGCGCACTTTGGCTGAACATGGAAGTAAAAA 2079
DB 7183 TAGATATGACAAAAGAAATGCTGAGAGGTGCGCCACCTGCTTCACATGAGAGTTAAGA 7242
QY 2080 TTATGATGAGTATATGATATCAAGCACTTACTTGTGTGGGAGATTTATCTGTGTG 2139
DB 7243 TCATGACGAGATCATGAGTGAAGAGACCACTTACTTGTGCGGAGATTTATCTTGAAG 7202
QY 2140 ACCAGATTAACGAGCAGCTGCAAGTCCGACACCTCTTAATAAGGCTTTTAAAGTTG 2199
DB 7203 ATTCTGTTACTTCCACAGCGTGTGCGGCGGACCCCTGAAAGGCTGTTTAAAGTTG 7262

Dh 9523 GGAGACTAGGGGCAAAACCCGGACCAACCACTGATGATGTCGGAAGAAGCGTCAAAA 9582
Qy 4383 ATTTTCTGTTGGGGGCGAAGAGGGCTGAGTACGATGAGGGGTAACCATGAACCACTGAG 4442
Db 9583 ACTTCACCGCTGACCGAGATGAGCTGGAAATACATATGAGGAATATGAGCCACTGAGGG 9642
Qy 4443 TCTGGGCGCAGAGTCCGACCAAGCGACCCACATGATGAGCGGCATGATGATCAATCC 4502
Db 9643 TCTATGCCCAAGATTCAGACACAGGAGACCTCAAGATGAGCCACAGAAATATATACAGC 9702
Qy 4503 ACTATATCATCGGATTCAGATTCATGATGATGATGATGATGATGATGATGATGATG 4562
Db 9703 ATTAATCACTATCGCATCTGATGATCAACATCTTATGCGCTGATGATGATGATGATG 9762
Qy 4563 TCCGATGAGGACATGATCATGACAGCTTGCATGCGCAAGCAAGAGAGCTGCTGA 4622
Db 9763 TGATGATTTGGGTAACCGTTGCAATGATGATGATGATGATGATGATGATGATGATG 9822
Qy 4623 CGCCATAGCGGCTTGCACCGAACCGCAACGATACCCACAGATTAAGCGGTTTGTCTGA 4682
Db 9823 CGCCATAGCGGCTTGCACCGAACCGCAACGATACCCACAGATTAAGCGGTTTGTCTGA 9882
Qy 4683 TTGCGCCAAACCAACGCTGAAACATTTGAGAACTTTGAACCATCTGTGTTAAACAAC 4742
Db 9883 TTAGTTCGCGCAATGCTGAAACGTTCAACGACCATGATTAATCTGTGTTAAACAAC 9942
Qy 4743 AACGTTTCTGTGGGACAGTTGATGATCTCTGAGGACGCTGTTATCTGTCTGCT 4802
Db 9943 AGCGTTTCTGTGGGACAGTTGATGATCTCTGAGGACGCTGTTATCTGTCTGCT 10002
Qy 4803 GCTTTTCAATGCTGATGCTTTTATTTATTTGTTGAGGCGTCTGCTGGGAGGATGACG 4862
Db 10003 GCTGCTCTGCTGCTGCTTTTATTTATTTGTTGAGGCGTCTGCTGGGAGGATGACG 10062
Qy 4863 CTTTGAACATGCAACCATCTGCAAAATGTTCCGGGATCCGATTAAGCGCTTGTG 4922
Db 10063 CTTTGAACATGCAACCATCTGCAAAATGTTCCGGGATCCGATTAAGCGCTTGTG 10122
Qy 4923 AACGCGAGGTTAGCGGCACTTAACCTGAGATCAGGATGATGATGATGATGATGATG 4982
Db 10123 AAAAGGCGAGGTTAGCGGCACTTAACCTGAGATCAGGATGATGATGATGATGATGATG 10182
Qy 4983 CTTTGAACATGCAACCATCTGCAAAATGTTCCGGGATCCGATTAAGCGCTTGTG 5042
Db 10183 CTTTGAACATGCAACCATCTGCAAAATGTTCCGGGATCCGATTAAGCGCTTGTG 10242
Qy 5043 TTAATGCTGCGGCTCTGATGAGCAAGGATCTTAAGGCGGATTAACATGCGCG 5102
Db 10243 TCAATGCTGCGGCTCTGATGAGCAAGGATCTTAAGGCGGATTAACATGCGCG 10302
Qy 5103 TTTTGGCGGTTGATCCCTTCAATGTTGGGAGGCGCAATGCTTCTGATGAGTGA 5162
Db 10303 TCTTGGGAGGTTGATCCCTTCAATGTTGGGAGGCGCAATGCTTCTGATGAGTGA 10362
Qy 5163 ACACACATGATGAGGCGTACGTCGATGCTGCTCAGATGATGATGATGATGATGATG 5222
Db 10363 ACACACATGATGAGGCGTACGTCGATGCTGCTCAGATGATGATGATGATGATGATG 10422
Qy 5223 TCGCATTAAGATTCACAGCTGCTGAAAGTGGCGCTGATGATGATGATGATGATG 5282
Db 10423 AGGCGATTAAGATTCACAGCTGCTGAAAGTGGCGCTGATGATGATGATGATGATG 10482
Qy 5283 CCACCGGCGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5342
Db 10483 CTACCACTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 10542
Qy 5343 AGGTATGACAGGCGCATATCAGCGCTTTTCAACCTTTGACCATTAAGGTCCTATCA 5402
Db 10543 AAGTATGACAGGCGCATATCAGCGCTTTTCAACCTTTGACCATTAAGGTCCTATCA 10602
Qy 5403 GAAAGGCGCTTTTATCAATCACTTCTGATGATGATGATGATGATGATGATGATGATG 5462
Db 10603 ATCCGCGCTGTTATCAATCACTTCTGATGATGATGATGATGATGATGATGATGATG 10662

Qy 5463 TCGCGATATTCAAGCATCTCGCTTGAATGCTACAGACATAGTACCCGACTGACATAC 5522
Db 10663 TTGAGACATTTCAAGTACTCTCTTGAATGCTACAGACATAGTACCCGACTGACATAC 10722
Qy 5523 GGTCTGTAAGCTTCTGTCAGAACATCCATGCTGCTTCAACCAAGCATTAAGGCT 5582
Db 10723 GGTCTGTAAGCTTCTGTCAGAACATCCATGCTGCTTCAACCAAGCATTAAGGCT 10782
Qy 5583 ATGAATGTAAGAACATCACTGAGACGACCCCTGCAAGAACATCACTGAGATGTA 5642
Db 10783 ATGAATGTAAGAACATCACTGAGACGACCCCTGCAAGAACATCACTGAGATGTA 10842
Qy 5643 AAATTAAGTGAAGCTCTGCGAGCTTCAATCTGCTTCAAGGCAATCTTATCTGA 5702
Db 10843 AGATTAAGTGAAGCTCTGCGAGCTTCAATCTGCTTCAAGGCAATCTTATCTGA 10902
Qy 5703 TTGACATCTCTGATGACGCTTTTGTAGATCATCAAGATCAACCAATTTAGAACTTA 5762
Db 10903 TTGACATCTCTGATGACGCTTTTGTAGATCATCAAGATCAACCAATTTAGAACTTA 10962
Qy 5763 GCTGCAAGTGAAGACATGATTTATTTGAGACATTTGATGATGATGATGATGATGATG 5822
Db 10963 AATGTAAGTGAAGACATGATTTATTTGAGACATTTGATGATGATGATGATGATGATG 11022
Qy 5823 ACAAGCTGACAGGAGGACATTTGTCAGTTCATCTCCATCCACAGACAGCTGTTTGA 5882
Db 11023 ATGATTCGACCGCGAAGGATCAATGCGGCTTCAATTTGATGATGATGATGATGATGATG 11082
Qy 5883 AGAAGCGACACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5942
Db 11083 AGAAGCGACACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 11142
Qy 5943 GCCCAGACGAATTTATTTATTTGCTATGCGGCAAGATCACTGCAATGCTGAAT 6002
Db 11143 GTCACAGGGAATTTATTTATTTGCTATGCGGCAAGATCACTGCAATGCTGAAT 11202
Qy 6003 GTAAACACCGGCGACACATTAATGAGAACCAATTAAGTGAACCAAGAAATTCAGG 6062
Db 11203 GTAAACACCGGCGACACATTAATGAGAACCAATTAAGTGAACCAAGAAATTCAGG 11262
Qy 6063 CGGCGATTTCAAAACATTTGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6122
Db 11263 CGGCGATTTCAAAACATTTGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 11322
Qy 6123 TCAATGTTGAGACATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 6182
Db 11323 TCAATGTTGAGACATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 11382
Qy 6183 GACTGACGCGGACACTGA-CATAGCGGTAAACTGATGATGATGATGATGATGATGATG 6241
Db 11383 GACTGACGCGGACACTGA-CATAGCGGTAAACTGATGATGATGATGATGATGATGATG 11442
Qy 6242 GCATTAATGCAAGCGGCTGCT 6261
Db 11443 GCATTAATGCAAGCGGCTGCT 11462

RESULT 5
AAV3474
ID AAV3474 standard; cDNA; 11717 BP.
XX
XX AAV3474;
DT 10-DEC-1998 (first entry)
XX
XX Girdwood S.A.virus cDNA sequence;
XX
XX Bone marrow cell; alphavirus; Girdwood S.A.virus; cytokine; capsid;
XX reverse transcriptase-PCR; nsP1; nsP2; nsP3; E3; E2; E1;
XX growth hormone; growth factor; interleukin; chemokine; enzyme;
XX ribozyme; antisense oligonucleotide; ss.

OS Girdwood S.A. virus.
XX Key Location/Qualifiers
FH 5'UTR 1..59
FT /tag= a
FT 60..7616
FT CDS /tag= b
FT /product= "Non-structural polypeptide"
FT /note= "Internal stop codon present at nucleotides 5763-5765"
FT 7662..11399
FT /tag= c
FT /product= "Structural polypeptide"
FT 11400..11717
FT 3'UTR /tag= d
XX W09836779-A2.
XX 27-AUG-1998.
XX 18-FEB-1998; 98MO-US02945.
XX 19-FEB-1997; 97US-0801263.
XX (UNNC-) UNIV NORTH CAROLINA.
XX Davis NL, Johnston RE, Simpson DA;
XX WPI, 1998-495361/42.
XX P-PDB; AAM70462, AAM70463, AAM70473, AAM70474, AAM70475, AAM70476,
XX AAM70477, AAM70478, AAM70479.
XX
XX Expression of heterologous RNA in bone marrow cells - using a
XX recombinant alphavirus comprising a promoter operable in bone marrow
XX cells operably associated with heterologous RNA.
XX
XX Example 2; Fig 3A-3C; 689p; English.
XX
XX The invention provides a method for introducing and expressing
XX heterologous RNA in bone marrow cells using alphavirus vectors. The
XX Girdwood S.A. virus is an alphavirus whose cDNA sequence was determined
XX from uncoded reverse transcriptase-PCR reaction fragments amplified
XX from the virion RNA. The Girdwood S.A. virus cDNA encodes nonstructural
XX and structural polypeptides. The nonstructural polypeptide is,
XX presumably, post-translationally modified into four different protein
XX products, namely nsP1 (AAM70462), nsP2 (AAM70463), nsP3 (AAM70473), nsP4
XX (AAM70474) proteins. The structural polypeptide is, presumably, post-
XX translationally modified into five different protein products, namely
XX capsid (AAM70475), E3 (AAM70476), E2 (AAM70477), 6K (AAM70478), E1
XX (AAM70479) proteins. The Girdwood S.A. virus cDNA was used in the method
XX of the invention. The inventors claim the transformed bone marrow cells
XX are useful for expressing a protein or peptide suitable for protecting
XX the subject against a disease such as a microbial, bacterial, protozoal,
XX parasitic or viral disease. The transformed bone marrow cells are also
XX claimed to be useful for expressing proteins and peptides such as
XX hormones, growth hormones, growth factors, interleukins, cytokines,
XX chemokines, enzymes, ribozymes or antisense oligonucleotides.
XX
XX Sequence 11717 BP; 3297 A; 3130 C; 2927 G; 2360 T; 3 other;
SQ

Query Match 36.8%; Score 2386.4; DB 19; Length 11717;
Best Local Similarity 64.5%; Pred. No. 0;
Matches 3745; Conservative 0; Mismatches 1991; Indels 74; Gaps 9;

QY 580 ACTGACGATGACGCGGAGCGTATTTCTCATCGGAAACAGGCCAAGGTCACTTC 639
DB 5761 ACTGACTAACCGGGGTAGGTGGTACATATTTTCAGCAGCAGAGCCCTGGGCACTTGC 5820
QY 640 AACGAAATTCAGTCAATGTAATCAAGAACTTATTTGATGGGCGGCTTCATG 699
DB 5821 AATGAGAGTCCGTTTGCAAGATGAGTACGAAACGACCTTGGAGGCAATGTTCTGG 5880
QY 700 AGAAGTATTACGCCCGGCTCGATCTCGAAGAAGAAATGTTACGAAGAATACTGC 759

DB 5881 AAAGATCTACGCCCGGCTGCTCGACAGCTGGAAGAAAGAGCAAGCTCAAACTCAGGTACC 5940
QY 760 AATTATGCGCTCTGGAAGAAATAGAACAGTATCAATCAAGAAAGTAAATATGCA 819
DB 5941 AGATGATGCGCCAGCAGAACCCAAACAAAGCAGGTACCACTCTAGAAATAGAAATCAGA 6000
QY 820 AAGCAATTAACGCGAGGAGCACTATTTCTGATTTGGGCAATATCTATCATCAGAAATGA 879
DB 6001 AAGCATTAACCATGAGGAGCTGCTTTAGGGCTACAGCTGTATCTCTGACAG--- 6057
QY 880 ATCTGTGAGTGTATACAGATCAATTAATCTGTACCAATCTACTCTCAACGATTAATTA 939
DB 6058 ATCAGCGGAATGCTATTAAGATCACTACCGGAACATCGATATTCAGAGGTGACGG 6117
QY 940 ACAGGTTTACATCTGCAAGAGTGGGCTTAAACGTGCACTTACTTATTCAGAGAAAT 999
DB 6118 CGAACTACTCTGACCCAAAGTTTCTGTAGCTGTGTTGCAACAACTATCTGATGAGAAAT 6177
QY 1000 ACCCTACAGTACCGGCTTATTTGTAACAGATGAATAGATGCGTATCTGACATGTTGG 1059
DB 6178 ACCGACGATGATCTTATCAATCACGACGAGTACGATGCTTACTTGATGTTAG 6237
QY 1060 ACGGCGATCGTGTCTAGATACAGCCACTTTTGTCCGGCTAAACCTGAGAACTACC 1119
DB 6238 ACGGACAGTGGCTTGTCTGATATCTGCAACTTTTGTCCGGCAAGCTTGAAGTTACC 6297
QY 1120 CAAAGACATAGCTATTTGCAAGCAGATTAAGATACGCGCTCCATGCTTATACGA 1179
DB 6298 CGAAAGACAGAGATATGAGACCCCAACCTCCAGTCCGCTTCACTACAGATGACGA 6357
QY 1180 ATACATTAACAAATGATATGCTGCTGAGTACTTAAAGAAATGCAAGCTTCCCAAAATGC 1239
DB 6358 ACAGCTTCAAAAGCTGCTCACTTGTCCGACTTAAAGAACTGCAAGCTCACAAATGC 6417
QY 1240 GAGATTAACCTGTCTTGAATTCGCGGCTTAAATGTTGATTTGTTCAAGAAATACGAT 1299
DB 6418 GTGAATTCGCAACCTGAGACTCAGAGCATTTCAAGTTGATGTTGAAATATGATCAT 6477
QY 1300 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1359
DB 6478 GTAAATGAGATATTTGGAGAGATTTGCGCAAAAGCCAAATGATGATGATGATGATG 6537
QY 1360 TTACGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1419
DB 6538 TTACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6597
QY 1420 AATAATCTAAACCGTGTGAGAGATATCAATGATGATGATGATGATGATGATGATGATG 1479
DB 6598 AATAATTTGCTCCATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 6657
QY 1480 ATGTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1539
DB 6658 ACGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6717
QY 1540 AGGCTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1599
DB 6718 AAGCGCAGAAACCCCTGCGACGCTTACCTGTGCGGATCACCGGAGTTAGTGCACA 6777
QY 1600 GACTGAATGCGGCTCTTCTGCAATATCAATGATGATGATGATGATGATGATGATGATGATG 1659
DB 6778 GCGTTACAGCGCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 6837
QY 6838 TTGATGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6897
DB 1720 CGTGCTGATTAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1779
QY 6898 CCTGCTGACAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 6957
QY 1780 ACTTAGTGTGACCAACCGCTTGAATTTGATGATGATGATGATGATGATGATGATGATGAT 1839

Dh 6958 ACTGGGTGTGAGCAACCACTACTGACTGATCGAGTGGCGCTTTGGAGAAATATCAT 7017
Qy 1840 CTGTGCACTTACCTTACAGAGGTTAAATTGGTGCATATGAAATCCGGTATGT 1899
Db 7018 CCACCCATCTGCGCAGGGAATCCGTTCAAAATTCGGGGCGATATGAAATCCGGAATGT 7077
Qy 1900 TCTTAAAGCTTTGTGTCAACACACTAGTCAATATCATGATTTGTAGCAGAGTACTACGTTG 1959
Db 7078 TCTTCAAGCTTTGTGTCAACACACTAGTCAATATGATTTGTAGCAGAGTACTACGTTG 7137
Qy 1960 AACGGTTAACCACTGAGCGTGGCGGCTCTATCGGCGAGATPACATAGTGTG 2019
Db 7138 AGCGGCTTAAACGTTCCAAATGTGACGATTTATCGGCGAGCAACATCATACGAG 7197
Qy 2020 TCGTCTCCGACACCTTGAATGGCGGAGATGCGGCACCTTGGCTGAACATGGAATAAAA 2079
Db 7198 TAGATCTGACAAAGAAATGGCTGAGAGTGGCACCTGGCTCAACATGGAAGTTAGA 7257
Qy 2080 TTATTTGATGAGTATTGTTATTCAAAGCACTTACTTGTGGGGGATTTATCCCTGTGG 2139
Db 7258 TCATTTGAGCGATCATCGGCGAGAGACCGCTTACTTGTGGGTGATTCATCTTGCAG 7317
Qy 2140 ACCAGATPACAGGACAGCGCTGACAGTGGCAGACCTCTAAAAAGCGTTTAAAGCTTG 2199
Db 7318 ATTGGTTACTCTCCACAGCGCTGTGCGGTGGCGGACCTTGAAAGGCTGTTAAGTTGG 7377
Qy 2200 GAAACCATTTGGCCGTGATATACCAAGACCTGCGACCGCGCGCGGCACTGATGATG 2259
Db 7378 GTAAACCGCTCCAGCGACGACGACGAGCAAGAAAGCAAGACGCGCTGTGATGATG 7437
Qy 2260 AAGCAATGCGATGAGAAAGAAATTTGAAATTACGAGCGATTTAGTAAGCGCTTAGAATCCA 2319
Db 7438 AAACAAAGGCGTGTGTTAAGATAGTATACAGACACTTAGCAGTGGCGCTGCAATC 7497
Qy 2320 GATACGAGATCATCTGGCAGCGCTGATCATCGCTCTGTGTCACGTTAGCCGAAACG 2379
Db 7498 GGTATGAGTGTGAGCAACATCACCTGTCTGTGATGAGAACTTTTCCGAGACA 7557
Qy 2380 TTAGAACTTCAAGATTAAGAGGAGGCCAATCACCTCTTAGCGCTGACCTTAATAG 2439
Db 7558 AAAGAGATTTCAAGCACTACAGAGGGAATTAAGCATCTCTACGGTGTCTTAATAGT 7617
Qy 2440 TGAGTATGTA-----GACACGACCTACCCGCGGAGAAATGTTTCATPACC 2486
Db 7618 CAGATATGACATTTTCACTCTGATTAATACCAACACCAACCATGATGAGATTTCT 7677
Qy 2487 CTCAGCTGAATTTTCAACCACTTACCTTACAAATTCGATGCTTACCGAGATCAACC 2546
Db 7678 TTATCATGTGTGCGCGCGCGCTTCCGCGCCCACTGCGATGTGAGCGCGGAGAA 7737
Qy 2547 CTCTAGCGCGCGCTGAGGCGCGTTTGGCCCCCGCTGGCTGCTCAATTCGAATCTTA 2606
Db 7738 GGAGGACGCGCGCGCGCGATGCTTCCCGCAATGGGCTGGCTTCCCAATTCAGCACTGA 7797
Qy 2607 GGAGGTGATGATGCACTTGAATTTCAACA-----ACGATCACTTAATCCGCGCAG 2660
Db 7798 CCAAGCGCTGAGTGGCTTACTGATTTGAGCAGGCACTAGACCTCAAAACCCACGCGCAC 7857
Qy 2661 GTCCACCGCCAAAGAGAAAGAGTGTCTCTAAAGCCAAACCTTACTCAGCTTAAAAA 2720
Db 7858 GCCCGCGCGCGCGCAGAAAGAGGCGCCAAAGCAACCAAGAGCGGAACCA 7917
Qy 2721 AGAAGCAGCAAGCCAAAGAGAGCAAAAC-----GCAAGCTTAACAGGGAAGAGCAAGTA 2777
Db 7918 AAACAGAGAAAGAAAGAAAGAAAGCACTGCAAAACCAAAACCGGAAAGAGCAAGTA 7977
Qy 2778 TGTATATGAAGTTGAGTGGAGCAAGACATTTCCGATC---ATGCTGAACGCGCAAGTA 2834
Db 7978 TGGCACTCAAGTTGAGGCGCAGACAGCTGTTGAGCTCAAAATGAGAGAGCAAGATGTCA 8037
Qy 2835 ATGATATGCTGTGCTGTGCGAGAAAGCTGATGAAACCACTCAAGCTTGAAGAAAA 2894
Db 8038 TCGGGACGCACTGCGCATGAGAAAGAAAGTAATGAAACCACTCAAGTGAAGAACTA 8097

Qy 2895 TTGATATGAGCAATTAGCGCGCGTGAATTTGAAGAGGCTAGCATGTACGATTGAGT 2954
Db 8098 TTGACACCTGTGTCTATCAAAAGCTCAAAATTCACAGATGTGACGATACGATGAGT 8157
Qy 2955 ACCGCGAGTGTTCACAGAAATGAGAAATCAGACAGCTGCAAGTACACCGAGCAAAACAC 3014
Db 8158 TCGCAGAGTTCCGGGTCAACATGAGAAAGTGAAGGCTTCACTTACACGAGGAACACCTG 8217
Qy 3015 CGGCTTCTTCAACTGACACAGCGCGCACTCCAGTATGAGAAATGGAGATTTACCGTAC 3074
Db 8218 AAGGTTTTCAACTGCGCACAGGAGCGGTGAGATTAAGTGAAGTTTATCCATCC 8277
Qy 3075 CGAGAGAGTGGCGCGGAAAGGAGAGCGGAAAGACCGGATCTGAGCAACAGAGAGAG 3134
Db 8278 CCGCGAGATGAGAGGAGAGAGACAGTGTGTGTCCGATTTATGATTACTCAGCGCGG 8337
Qy 3135 TTGTGCTATTTGTTAGAGGTGCAAAATGAGGACGCGTACGCGCTTTCAGTGTCA 3194
Db 8338 TTGTGCGATAGTCTCGAGGGGCTGATGAGGGAACAAAGACTGCCCTTTCGGTGTCA 8397
Qy 3195 CTGGAACCAAGAAAGGCTGACATTAAGGATACCCCGAAAGTTTGAACCGTGT--- 3251
Db 8398 CTGGAATAGCAAGGAGAGACATCAAGCAACCCCGAAAGGAGACAGAAAGTGTCTG 8457
Qy 3252 -----CACTAGTTACAGCGCTATGCGTGTCTTGAATGTACAGTTCCATGCGACAAC 3305
Db 8458 CAGCACCACTGTGTACGCGCATATGTCTTGTGAAACGTAAGCTTCCATGCAATGGCC 8517
Qy 3306 CACCGTGTCTATTCACTGACCGCAGAACCAACTGACGTGTGAAAGAAACGTG 3365
Db 8518 CGCCCATGTCTACACCCCGGAACATCCAGAGCTTGAATCTTGAAGAAACGTGA 8577
Qy 3366 ACAATCAATTTACAGACGCTGTGAGAAAGCTTTGAAT---GTCCATCACGCGGC 3422
Db 8578 ACCAGAGGCTTACAGACCTGTCTCAACGCAATTTGCGGTGCGATCTCCGCGAGAA 8637
Qy 3423 CCAAGGACATTTACGATGACCTTCACTGACAGTCCCTTACCTGCGGTTTGCCTG 3482
Db 8638 GCAAAAGAGCGTCTACGAGACTTTTACCTTGAACAGCCGCTTGTGGCAGACATGCTGT 8697
Qy 3483 ATTGCAACACTGACCGCGTGTTCAGCCCAATAAAAATGAGAAAGTGTGGAGCAAT 3542
Db 8698 ACTGTACCATATGAAACGTTGCTTACCGCATTAAGATCGAGAGGTCTGGGATGAG 8757
Qy 3543 CTGATGATGATGATTAAGATCCAGGTCTGCGCAATTCGGCTACATCAGCAGGCA 3602
Db 8758 CGGACGCAACACCATACGATACAGACTTCCGCGCGATTTGATACGACCAAGCGGAG 8817
Qy 3603 CTGCGATGTCACCAAAATTCGTTTACATGTCTTTCGACAGACCATGATCATCAGGAG 3662
Db 8818 CAGCAAGCTCAAAATTAAGTACCGCTACATGTGCTCGACAGATCATACCTCAAGAG 8877
Qy 3663 ACAATATGAGAAATTAAGTATGACGACATCTGACCCGCTGCTTGTGGCCACAAG 3722
Db 8878 GCACTATGATGACATCAAGATCAGCACTCAGGACCGGTGAGAGGCTTAGCTACAAAG 8937
Qy 3723 GGTACTTCTGTAGCTCAATGTCTTCAGAGTGAAGTAAACCGTCAAGTACAGGCG 3782
Db 8938 GATACCTTCTCTCGCGAAGTGTCTTCAGGAGGACAGCTTAAAGGTTAGATGCGAGTA 8997
Qy 3783 GAGACTGAGATTTCACTGACCGGTGAGAAAGATCAGAGAGAAATTTGTGCGTGAAG 3842
Db 8998 GCAACTAGCAAGTCACTGACATGAGCGCGCAGAGATTAACCAAAATTTGTGGAGCGG 9057
Qy 3843 AGAGTACTTGTTCACACCGCTCATAGAAAGCTGTGAAGTCCACGTTACGATCAT 3902
Db 9058 AAAAATATGACCTTACCTCCGTTACGTTAAGAAATCTTTCACAGTGTACGACGCTC 9117
Qy 3903 TGAAGAGAGCTGTGCGGTATCATACATGACAGGCGGACCAACAGCGTATTAAGT 3962
Db 9118 TGAAGAAACAAACCGCGGTATCATCATATGACAGGCGGAGACGCAAGCTTATACGT 9177

QY 3963 CCTATCTGAGAGAGCGTCAGCGGAAGTGTACATTAAACACCTTCTGGCAAGACGTCA 4022
Db 9178 CCTATCTGAGAGAAATCATACGGGAAAGTCTACGCCAAGCCACATCCGAAAGAACATTA 9237
QY 4023 CCTAGGAATGTAGTGTGCGCACTACAGACAGATATCTGTAGACAGCGCAAGAAATGA 4082
Db 9238 CTTAGAGAGTCAAGTGGCGGATTCACAGACCGGTACCGTTACGACCCCTACCGAAATGA 9297
QY 4083 ACGGCTGCACTAAAGCAAAACAGTGCATTGCTACAAAGACGCAACCAAGAAATGGGTCT 4142
Db 9298 CGGGCTGACCGCCCATCAAGAGTGTGTGCTATTAAGGCCAACCAAGAAATGGGTCT 9357
QY 4143 TCAACTCCGCGGATCTTATTAGGCAACAGACCACTCAAGTGCAGGTAATTTGACATTC 4202
Db 9358 TCAATTTCCGCGGATCTTATGATCAGACATGCGACCAACGCGCCCAAGGAAATTTGACATTC 9417
QY 4203 CATTTCCGCTTGACACCGACAGTGTGCGGCTTCCGTTAGCTACACGCTTACAGTACGA 4262
Db 9418 CTTTCAAGCTGATCCGAGTACCTGTGATGTCCCTGTTCGCCACGCGCGGAAGTATAC 9477
QY 4263 AGTGTTCAAAGGATCACCTCTCACTGACTGCAATGGCAACATTTGTCACAAAGA 4322
Db 9478 ACGGCTTTAAACATCAGCTCTCAATTAGACACAGACCACTGACATTTGTCACCA 9537
QY 4323 GAAAAATGGGGGTGCGAGCAGACGCAACAGAAATGATTAAGGGTCTACATCAGGA 4382
Db 9538 GGAAGCTAGGGGCAATCCGGAACCAACTGATGATTCATCGGAAAGAGGTTAGAA 9597
QY 4383 ATTTTCTGTGGGGGAGAAAGGCTGGAGTACGTATGGGTGTAACCAATGACATCAG 4442
Db 9598 ACTTACCGCTCAGCAGAGATGCGCTGGAATCATATGGGGCAATCAAGAACGGAAGG 9657
QY 4443 TCTGGGCGGAGGTCGCGACGAGGAGCCCAATGATGCGCGGATGATCATCATC 4502
Db 9658 TCTATGCGCAAGTCTGCAAGAGAGACCTTCAAGATGCGCAACGAAATGATCAGC 9717
QY 4503 ACTATTTATCATGGGATCCAGTCTACACTGTATGTGTGTGTGTGTGTGTGTGTGT 4562
Db 9718 ATTATCTACATCGCCATCTGTGTACACCATCTTAGCCGTGCATCAGTGTGTGTGTGT 9777
QY 4563 TCTGT 4622
Db 9778 TGATATTTGT 9837
QY 4623 CGCCATGACGCGCTTGACCGAAGCGAAGCGATCCCAAGCATTTAGGGTTTTGTGTGT 4682
Db 9838 CGCCATGATGCGCTTGCGCCCAAAATGCCGTGATTCAACTTGTGTGTGTGTGTGTGT 9897
QY 4683 TTTCGCAACCAACGCTGAAACATTTGAGAACTTTGAACCATCTGTGTGTGTGTGTGT 4742
Db 9898 TTAGGTCCGCTTAATGCTGAAACATTTACCGAAGCATGATGTTACCTATGTGTGTGTGT 9957
QY 4743 AACGTTTCTGTGGGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4802
Db 9958 AGCCATTTCTGTGGGTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10017
QY 4803 GCTTTTCAATGCTGACGCTTTTATTTATGTTGTGTGTGTGTGTGTGTGTGTGTGTGT 4862
Db 10018 GTTGTCTATGCTGTGCTTTTATTTATGTTGTGTGTGTGTGTGTGTGTGTGTGTGT 10077
QY 4863 CTTTGGAACATGCAACATGTCGCAAAATGTTCCGGGAGATCCCGTATTAAGGGTGTGTG 4922
Db 10078 CCTAGGAACATGCAACATGTTTCCAAATGTGCAAGATACCGATGTTAGGCACTTTGTG 10137
QY 4923 AAAGGCAAGTTTACGCGCACTTAACTGTGAGATCAAGGTGTGTGTGTGTGTGTGTGTGT 4982
Db 10138 AAAGGCAAGGTGTACCGCGCTCAATTGTGAGATTAATGTGTGTGTGTGTGTGTGTGT 10197
QY 4983 CTTCAACTAAGAGAGTACGATGCAATTTCCAGCACTCATCTTTTACCAAG 5042
Db 10198 CTTTCAACCAAGAGATCATCCTGTGAAATTCACCACTGTGTGTGTGTGTGTGTGTGT 10257
QY 5043 TTAAATGTGTGCGGTCTGTGATGCAAGGATCTCTCAAGGCGGATTCACATGCTGTG 5102

Db 10258 TCAATGTGTGCGGTCTGTGATGTACAGCCGCGCTTACAGCAAGTATACCTGCAAG 10317
QY 5103 TTTTGGGCGGTGTGATCCCTTTTCAATGTTGGGAGGCGCAATGCTTCTGTGACATGAGA 5162
Db 10318 TCTTTGGAGGGGTGTATCCCTTTCATGTGGGAGGAGCAACAAATTTTGGCAGATGAGA 10377
QY 5163 ACACAACTGAGTGGAGGCTGATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5222
Db 10378 ACAGCAATGATGATGAGCGTATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10437
QY 5223 TCGCACTAAAGTTTCAACAGCTGTGTGAAATGTGGGCTGTGTGTGTGTGTGTGTGTGT 5282
Db 10438 AGGCGATTAAGGTGATCTGTGCGGATGAAAGTGTGTGTGTGTGTGTGTGTGTGTGT 10497
QY 5283 CCAAGCGGACCTGTGATACGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5342
Db 10498 CTTACAGTTTCTTATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10557
QY 5343 AGGTATGACAGGCGGATGATGAGCGGCTTTTTCACCTTTGACCATTAAGGTGTGTGTGT 5402
Db 10558 AAGTATGATGTGACCAATTTTCAAGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10617
QY 5403 GAAAGGGGCTGT 5462
Db 10618 ATCGGCGCTGT 10677
QY 5463 TCGGCGATTTTCAAGCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5522
Db 10678 TTTGAGATTTTCAAGCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10737
QY 5523 GCGTGTGAGGCTTGT 5582
Db 10738 GACTATGATGCTGT 10797
QY 5583 ATGAATGTGTGAAGAACTGAGGACGACCCCTGTGCAAGAAACAGACATTTGTGTGTGT 5642
Db 10798 TCGAGATGTGTGAAGAACTGAGGACGACCCCTGTGCAAGAAACAGACATTTGTGTGT 10857
QY 5643 AAATTTGAATGAGGCTGT 5702
Db 10858 AGATTTGAGTCAATCCGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10917
QY 5703 TTGACATCCCTGTGATCCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5762
Db 10918 TCGACATCCGATCGCTGT 10977
QY 5763 GCTGCAAGTGTGACGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5822
Db 10978 AATGTGATGTGATGT 11037
QY 5823 ACAAGTGTGACGAGGAGGACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5882
Db 11038 ATGTATCCGAGCGGCAAGAAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11097
QY 5883 AGGAAGGACCAACATGT 5942
Db 11098 AAGATGTGACAGTTATGT 11157
QY 5943 GCCCAAGCAAAATTTTATGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6002
Db 11158 GCCCAAGGCAAACTTTTATGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11217
QY 6003 GTAAACCAAGCGGACCACTTAATTTGAGAAACCATTAAGTGTGTGTGTGTGTGTGTGT 6062
Db 11218 GCAAAACCAAGCGGACCACTTAATTTGAGAAACCATTAAGTGTGTGTGTGTGTGTGTGT 11277
QY 6063 CGGCAATTTTCAAAACATTTGT 6122
Db 11278 CGGCAATTTTCAAAACATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11337
QY 6123 TCAATTTGT 6182

Db 11338 TATTAAATATAGACTTATGATTTTCTTCAGCATGATCTGACTAGCACAGCAAGAT 11397
Qy 6183 GACTGAGGCGGAGACACGA-CATAGCGGTAAACCTGATCTCTCGAGGAGACGGTGT 6241
Db 11398 GACCGCTACCGCCCAATGACCCGACACAGCAAACTCGATGATCTCGAGGAACTGATGT 11457
Qy 6242 GCATATGCGCACGCGCGCT-----TGACACT 6268
Db 11458 GCATATGCGCATGAGGCTGTATATTAAGTCCCGCTTACCGGGGCAATATAGCAACACC 11517-
Qy 6269 AAAAAGTGAATATTTCCGAGGAGACAGTCAATATGCTGTGATGT 6318
Db 11518 AAAAAGTGAATATTTCCGAGGAGACAGTCAATATGCTGTGCGACATGT 11567
RESULT 6
AAC64506 standard; DNA; 11703 BP.
AAC64506;
AC AAC64506;
XX 14-FEB-2001 (first entry)
DT 14-FEB-2001 (first entry)
XX Alphavirus Sindchiron virus nucleotide sequence.
XX Alphavirus; Sindchiron virus; SindchironLP virus; immune response;
KM infection; human dendritic cell; immunostimulatory; cytostatic;
KM virucide; fungicide; antibacterial; antiparasitic; vaccine; cancer;
KM pathogen; antigen presenting cell; ds.
XX Alphavirus.
OS
XX WO20061772-A2.
PN 19-OCT-2000.
PD 14-APR-2000; 2000WO-US10722.
PF 14-APR-1999; 99US-0129498.
XX 09-AUG-1999; 99US-0148086.
PR 22-MAR-2000; 2000US-0191363.
XX (CHIR) CHIRON CORP.
PA
PI Polo JM, Dubensky TW, Frolov I, Gardner JP, Otten G, Barnett S;
PI Driver DA.
XX WPI; 2000-619231/59.
DR
XX New alphavirus that infects human dendritic cells for use in generating
PT an immune response to pathogenic agents such as bacteria, viruses,
PT fungi, parasites and cancer and for biological assays -
XX
XX Claim 9; Fig 2B; 83p; English.
PS
XX The present invention describes an isolated alphavirus (AV) which
CC infects human dendritic cells and is not of American Type Culture
CC Collection (ATCC) number VR-2526. AAC64506 and AAC64507 represent the
CC nucleotide sequence of the specifically claimed Sindchiron virus and
CC SindchironLP virus. The new AVs have immunostimulatory, cytostatic,
CC virucide, fungicide, antibacterial and antiparasitic activities and
CC can be used in vaccines. The AVs are used to infect dendritic cells,
CC preferably human cells. A heterologous sequence can be introduced and
CC expressed in human macrophages or antigen presenting cells in vivo and
CC in vitro, for use in biological assays. The AV-based vector systems
CC are used to generate an immune response to cancer or a pathogenic
CC agent, such as, bacteria, fungi, parasites or viruses. The AV can be
CC used to infect human dendritic cells, macrophages or antigen
CC presenting cells that previously could not be infected using an AV or
CC AV variant. The AV vectors are targeted directly to antigen presenting
CC cells.
XX
SQ Sequence 11703 BP; 3310 A; 3049 C; 2910 G; 2434 T; 0 other;

Query Match 36.8%; Score 2383.6; DB 21; Length 11703;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3695; Conservative 0; Mismatches 1984; Indels 41; Gaps 8;
Qy 580 ACTGACGGTATGAAGCGGAGCGGTATATTTCTCATCGGAAGCAAGGCTACCTTC 639
Db 5746 ACTGACTAACCGGGGTGTGTGGTATCATATTTTCAGCGGACACAGGCCCTGGACACTTGC 5805
Qy 640 AACGAAATATAGTACGTCATATGTAATCAAAACCTATATGATCGGGCCGTCCATG 699
Db 5806 AAAAGAAAGTCGGTTCGCAAAACAGTTTACAAACCGACCTTGGAGCGGAATGTCTTG 5865
Qy 700 AGAAGTATTAAGCGCGCGCGCTGATCTGAAAGAGAAATGTTACAGAAAGAACTGC 759
Db 5866 AAAAGATTCATGCGCCCGGTCTGACACGTCGAAAGAGAAACAACTCAAGTACC 5925
Qy 760 AATTATGCGCCTGTGAAGAAATAGAAAGCATCAATCAGCAAAAAGTAAATATGA 819
Db 5926 AGATGATGCCCGCAAGCCAAACAAAGTATGATCAGATCTCGTAAAGTAAATATGAGA 5985
Qy 820 AAGCAATTACAGCGGAGCGACTATTTCGATTTGGGACATATCTATCATCAGAACTGA 879
Db 5986 AAGCATTAACCACTGAGCGCACTACTGACAGACTACGACTGTAACTCTGCAACAG--- 6042
Qy 880 ATCTGTGAGTGTTCAGAGTCAATTAATCTGTACCAATCTACTGTCGAACGTTAATTA 939
Db 6043 ATCAAGCCAGAAATCTATTAATATCACTATTCGAAACCATTTGTACTCAGTAGGTACCG 6102
Qy 940 ACAAGTTTACATCTGCAGAGGTGCGGTTAAACGTGCAACTTATGTTTCAAGAAATT 999
Db 6103 CGAATCTACCTCGATCCACAGTCTGCTGTAGCTGTCTGTAACTATCTCATAGAACT 6162
Qy 1000 ACCCTACAGTACGCAATATTTATTAACAGATGAATACGATGCGTATCTGATAGTGG 1059
Db 6163 ATCCGACAGTAGATCTTATCAATTAAGTACAGATGATGATCTTCTGATATGTTAG 6222
Qy 1060 ACGGCGCATCTGCTGTCTGATATACAGCCACTTTTGTCCGGCTAAACTGAGAAGCTACC 1119
Db 6223 ACGGACAGTCCGCTGCTGCTGATATGTCGAAACCTTTCGCCGCTTAAGATTAC 6282
Qy 1120 CAAAGAAAGCATATTTTGCAGCCAGAGATTAAGATCAGCGTCCATGCTTATACAGA 1179
Db 6283 CGAAGAAACATGATATGATAGAGCCCGCAATATCCGAGTGGGTTCCATCAGCATCAGA 6342
Qy 1180 ATACATTAACAAATGATTTGGCTGTCAGCTACTTAAAGAAATTCGATCCCAATGTC 1239
Db 6343 ACAAGCTACAAATGATGCTCATTGCGCACTTAAAGAAATTCGATCCGAGATGC 6402
Qy 1240 GAGATTAACCTGCTTAGATTTCCGCGGCAATTTATGTTGATTTTCAAGAAATACGCAT 1299
Db 6403 GTGAATCTGCAACACTGACTGACGCACTTCAATGTCAATGCTTTGCAAAATATGCAAT 6462
Qy 1300 GCAATGATGATGATGCGGATACCTTTTCCGATTAACCTTATGCGCTTACTACAGAGAACG 1359
Db 6463 GTATATGACAGTATTTGGAGAGAGTTCGCTCGAAGCCAAATTAGATTAACACATGAGTTTG 6522
Qy 1360 TTACGCAATTTGGAACAAAGCTTAAAGGCGGCAACAGAGATTTGTTCCGAATATCTC 1419
Db 6523 TCACCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6582
Qy 1420 ATATATCTAAACCGTTGCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1479
Db 6583 ATATATTTGCTCCATTTGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 6642
Qy 1480 ATGTCAAAAGTTACTCCCGGACGAAACATACAGAGAGCGGCTTAAAGTGTATATTC 1539
Db 6643 ACGTGAAGTTTACACGAGGACGAAACACAGAAAGAAAGTGTATGATGATGATGATGATGAT 6702
Qy 1540 AGGCTGCAATCCCTTGTCTACCGCTTACCTTTGGCGGATCCATCCGGAATTTAGTCCGTA 1599
Db 6703 AAGCCCGAGAAACCTCGGCACTGCTTACTTATGCGGATTCACGCGGAATTTAGTCCGTA 6762

QY 1600 GACTGATGCGGTCTTTCGCCAAATATTCATATCTCTTGCGACATGTCAGCGGAAGATT 1659
Db 6763 GCGTTACGCGCGCTTGTCTTCCAAACATTCACACGCTTTTGAATGTGCGGGAGGATT 6822
QY 1660 TTGATGCGATTATTTCTGAACATTTTCCACACGCGGACCCAGTATTTGGAAACGCAATCG 1719
Db 6823 TTGATGCAATCATYACAGAACATTTTCAGCAAGGAGACCCGGTACTGAGAGCGGATATCG 6882
QY 1720 CGTCTGTTGATTAAGCGAAGACGAGCGCTATCGCCATTTCCGCTTTGATGATCTTGAAG 1779
Db 6883 CATCATTTGACAAAGCCAAAGCGACGCTATGCGCTTAACCGGTCTGATGATCTTGGAG 6942
QY 1780 ACTTAGGTGTCGACCAACCGCTCTTAGATTGATAGAGCGCGCTTCCGCAATATCAAT 1839
Db 6943 ACTTGGGTGTGATCAACCACTACTGACTGATGAGTGCCTTTGGAGAAATATCAT 7002
QY 1840 CTGTGCACCTACCTACAGGAACGAGTTTAAATTTGGTCCATGATGAAATCCGTAATG 1899
Db 7003 CCACCATCTACCTACGCGTACTCGTTTAAATTCGGGGCGATGATGAATCCGGAATGT 7062
QY 1900 TCTTAAGCTGTTTGTCAACACACTAGTCAATATCATGATTGCTAGCAGACTACGTTG 1959
Db 7063 TCTCAACATTTTGTCAACACAGTTTGAATGTGTTATGCGCAGCAGACTACGAG 7122
QY 1960 AACGGTTAACACGTCAGCGTGCAGCGCTCTATCGCGACGATTAACATAGTCATGTG 2019
Db 7123 AGCGCTTAAACGTCAGATGTGCAGCGCTCATTTGGGAGCGCAACATCATACATGAG 7182
QY 2020 TCGTCTCCGACACCTTGTATGCGGAGAGATGCGCACTTGGCTGAACATGGAATTAAGA 2079
Db 7183 TAGATCTGACAAAGAAATGCTGAGAGTGCACACCTGCTCAACATGAGGTAAAGA 7242
QY 2080 TTATGATGCACTTATGTTGATCAAGCACCTACTCTGTTGGGGGATTTATCCTGATG 2139
Db 7243 TCATGAGCGCACTATCGGTGAGACACACTTACTCTGCGCGGATTTATCTTGAAG 7302
QY 2140 ACCAGATTAACAGGACAGCCTGACAGTGCAGACCTCTTAAGAAAGGCTTTTAAAGTTG 2199
Db 7303 ATTGGTTACTTCCACAGGTGCGCGTGGCGGACCCCTGAAGAGGCTTTAAAGTTGG 7362
QY 2200 GAAAAACCTTTCAGTCCATGATATACCCAAAGACTGCGACCGCGCGGCACTGCATGATG 2259
Db 7363 GTAACCCCTCCACCCGACGACGAGCAAGCAAGAAAGACGCGCTGCTAGATG 7422
QY 2260 AAGCAATGCAATGGAACGAATTTGAATTAAGGAGATTGTAAGTGAAGCCGATGAATCA 2319
Db 7423 AAACAAAGCGGTGTTTGAATGATTAACAGCACTTATGACATGCGCGTGAAGCACC 7482
QY 2320 GATACGATCATATCTGCGAGCGCTGATCATCAGTCTCTGTCCAGTTAGCCGAAGCG 2379
Db 7483 GGTATGAGGTAGACAAATTAACCTGTCTACTGCGATTTGAGAACTTTGCGCAGAGCA 7542
QY 2380 TTAAAGACTTCAAGCAATAGAGGAGGAGCCCAATCACTCTTACGGCTGACCTTAATAG 2439
Db 7543 AAAGGCAATTCACACCATAGAGGGGAAATTAACATCTTACGCTGCTCTTAATAGT 7602
QY 2440 TGAGGTAGTACAGCAGCACTAC-----CCACGCGGAGAAATGTTTCCATAC 2486
Db 7603 CAGCATAGTACATTTCACTGTGCTAATATCTACACACCAACCACTGAATAGAGGATCT 7662
QY 2487 CTCAGCTAACCTTTCACCAAGTTTACCTTACCAATCCGATGGCTTACGAGATCCAAAC 2546
Db 7663 TTAAATCTCGGCGCGCGCGCTTCCGCGCGCCCACTGCGCATGTGAGGCGCGGAGAA 7722
QY 2547 CTCCTAGCGCGCTGAGGCGGCTTTCGCGCGCGCTGCTGCTCAAAATCCAGCAATCGA 2606
Db 7723 GAGGCGAGCGCGCGCGTGTGCTGCGCGCAACGCGCTGCTTCAAAATCCAGCAATCGA 7782
QY 2607 GGAGGTGATAGTAACTTGAATTTCAACAAACATCACTTAATCCGCGCGCAAGTCCA- 2665
Db 7783 CCACAGCGGTACGTGCTTATGATGAGGCAATAGACCTTAACCCCGACCTCCAC 7842
QY 2666 -----CCGCAAAAGAAAGAGTGTCTTAAGCCAAACCTTACGACTAAGTAAAGA 2720

Db 7843 GCCCGCACCGCGCCAGAAAGCAGCGCGCCCAAGCAACCAAGCCGAAGAAACCA 7902
QY 2721 AGAAGCAGAACCCAGAGAGCAAGAAC--GCAAGCTTAAATCCAGGAAAGCAACGTA 2777
Db 7903 AAAGCAGAAAGAAAGAAAGCAACTGCAAAACCCGAAGAAAGAGACGCCA 7962
QY 2778 TGTGATAGTGTGAGTCCGAGCAAGACATTTCCGATCATG---CTGAAGCGCCAAAGTGA 2834
Db 7963 TGGCATTAATTTGAGGCGCCAGATGTTTGCAGCTCAAGAACGAGACGAGATGTCA 8022
QY 2835 ATGATATGCTGCGTGTGCGAAGAGCTGATGAACCACTCCACGTTAAGGAAAA 2894
Db 8023 TCGGCGACGACTGCGCCATGGAAGAAAGTAAATTAACCTTGCACGTTGAAGAACCA 8082
QY 2895 TTGATTAAGACATTTAGCGCGCGTGAATTTGAAGAGGCTAGCATGATGATGAGT 2954
Db 8083 TCGACACCTGTGTCTATCAAGCTCAAAATTTACCAAGTGTCAAGCATATGATGAGT 8142
QY 2955 ACGGCGAGCTTCCCGAAGCATGAATCAGACAGCTGCAGTACACGACGCAAAACAC 3014
Db 8143 TCGCAGCTTCCAGTCAACATGAGAGTGAAGCATTTCACTACACAGTGAACCCCG 8202
QY 3015 CGGCTTCTACACTGCGACACGCGCGGCTCAAGTATGAGAAATGGAATTTACCGTAC 3074
Db 8203 AAGATTTATATATGCGCACACGAGCGGTGAGTATGAGAGTATGATTTACATCC 8262
QY 3075 CGAGAGAGTGGCGGGAAGAACCGACAGCGGAACCGATCTTGAACAACAGGCAAG 3134
Db 8263 CTGCGGAGTGAAGGCGAGAGAGACAGCGGTGTGCTCCATATGATGATTAATCTCCGTCGG 8322
QY 3135 TTGTGCTAATGTTCTAGAGAGTGAATGAGGAGCGAGCGGTACCGGCTTTCAGTGTCA 3194
Db 8323 TTGTGCGATAGTCTCGGTGAGTGAATGAAGAACGAACTGCTTTTCGCTGTCA 8382
QY 3195 CTGGAACAGAAAGGGTGAACATTAAGGATACCCCGGAAGTTCGAAACCGTGT--- 3251
Db 8383 CCGGAATATGTAAGGGAAGCAATTAAGACGACCCCGGAAGGGAACAGAGTGTGCG 8442
QY 3252 -----CACTAGTTACAGCGCTATGCGTCTTTGAAATGTCACTTCCATGCGCAAAAC 3305
Db 8443 CAGACCACTGTGTACAGGCAATGTGTTGCTCGGAAATGATGAGCTTCCATGCGACCC 8502
QY 3306 CACCGGTGTATTTACCTGACGCAAGAACGAACTGAGTGTGGAAGAACGTCG 3365
Db 8503 CGCCCATGTCTATACCGCGCAACCTTCCAGAGCCTTGAACATCTTGAAGAAACGTGA 8562
QY 3366 ACAATCCAAATTAAGACACGCTGTGAGAAAGCTTTGAAATGTCAATCAC---GCCGCG 3422
Db 8563 ACATGAGCGCTAGATACCTGCTCAATGCAATGCGGTGCGGATCGTCTGCGAAG 8622
QY 3423 CCAAAAGCAATTAACGATATCTTACATGACATGACCAAGTCTTACCTGGGATTTGCGCGT 3482
Db 8623 GCAAAAGAGGTCACTGACCACTTATCCTGACAGCGCCCTTACTTGGGCAATGCTGT 8682
QY 3483 ATTGACACACTCAACGCGGTGTTACGCCCAATTAATAATGAGAACGTGAGGAGCAAT 3542
Db 8683 ACTGCACATATGTAACCGTGTTCAGCCTGTTTGAATGAGAGAGGTCTGGGACGAG 8742
QY 3543 CTGATGATGATGATTAAGATTAAGATTCAGGCTTCCGCAATTCGCTTACATCAGGAGCA 3602
Db 8743 CGAGAGTAACACATAGCATACAGACTTCCGCGCAATTTGATGACCAAGAGCGAG 8802
QY 3603 CTGCGATGTCAACAAATTCGTTATCATGTCTTTTGCACACGACCATGACATCAAGAG 3662
Db 8803 CAGCAAGGCGCAAAAGATCCGCTATCATGTGCTTAAAGCAAGATCACACCTTTAAAGAG 8862
QY 3663 ACAGTAAGAGAAATAGATATGACCATGTGAACCTGCGGTCTTGGCCACAAAG 3722
Db 8863 GCACCATGATGACATCAAGATTAAGACCTTGAAGCGGTGAGAAAGTATGATCAAAAG 8922
QY 3723 GGTACTTCTGTGATCAATGTCTCTCAGGTGACAGTGAACGCTGATCAAGAGCG 3782

Dh 8923 GATACCTTCTCCGCAAAATGCCCTCCAGGGGACAGCGTAAAGCTTAGCATAGTAGTA 8982
Qy 3783 GAGATCTGAGAAATTCATGACCGGTGAGAAAAAGATCAGAGAGAGATTGTCCGTAGAG 3842
Db 8983 GCAACTCAGCAACGTCATGTACTGCTGGCCCGCAAGATTAACCAAAATTCGTGGAGCGG 9042
Qy 3843 AGAGATCTTGTCCCAACCCGTCATGGAAGCTGTAAAGTGCACGTTTACATCACT 3902
Db 9043 AAAAATATGATCTACCTCCCTTACAGGTAAAAAAATTCCTTGCACAGTGTACACCGTC 9102
Qy 3903 TGAAGAGACGTCTGCGGGTACATTAACATGCAAGCGCACGCGCTTAAAGT 3962
Db 9103 TGAAGAAACAACTGACGGCTACATCACTATGCAAGCGCCGGACCGCACTTATACAT 9162
Qy 3963 CCTATCTGAGAGAGCGTCAGCGCAAGGTACATTAAACCACTTTCGCAAGAACGCA 4022
Db 9163 CTATCTGGAAGAAATCATCAGGGAAAGTTTACGAAAGCCGCACTCGGAAAGAACATTA 9222
Qy 4023 CCTACGAATGTAAAGTGTGACACTACAGCAGGTATCTGAGACCGCAAGATGA 4082
Db 9223 CGTATGAGTGAAGTGGCGGCACTACAGACCAAGCCGTTTGCACCCGCAACCAATCA 9282
Qy 4083 ACGGCTGCACTAAAGCAAAACAGTGCATTGCTTCAAGAGCGCAACGAATGGGTCT 4142
Db 9283 CTGGTTGACCGCCATCAAGCAGTGGCTGCTTAAAGAGCGACCAACGAAGTGGGTCT 9342
Qy 4143 TCAACTGCGCGGATCTTATTAGGCACACAGCCACTCAGTGCMAAGTTAAATTCACATTC 4202
Db 9343 TCAACTCAACCGGACTTGTATCAGACATGACACACACCGCCCAAGGGAAATTTGATTTGC 9402
Qy 4203 CATTCGCGTGAACCGCAGTGTGCGGCTTCCGTTAGCTCACACGCGCTCAGTCAAG 4262
Db 9403 CTTTCAAGTTATCCCGAGTACCTGATGTCCTCTGTTGCCACGCGCGGATGTAAATAC 9462
Qy 4263 AGTGTCTAAAGGATCACTCCCTCACTGACTGCAATGCGAACAACTTGTGCAACGA 4322
Db 9463 ATGCTTTAAACATTCAGCTCCCAATTAGATACAGACCACTTGACATTTGCTCACACCA 9522
Qy 4323 GAAATTTGGGCTCGAGACACCGCAACGCAAGATTTACAGGGTCTACATCCAGGA 4382
Db 9523 GGAAGCTAAGGGGCAAAACCGGAAACCACTGAATGATCGTGGAAAGCGGTCAAGAA 9582
Qy 4383 ATTTTCTGTGGGCGAGAAAGGCTGAGTACGTATGGGTAAACATTAACAGTCAAG 4442
Db 9583 ACTTCAACGCTGACCGGAGATGGCTGTGAATCATATGGGAAATCATAGCAGTGAAGG 9642
Qy 4443 TCTGGGCGCAGAGTGGGACCAAGCGAACCAATGATGCGCGCATGATCATATCC 4502
Db 9643 TCTATGCCAAGAGTCAACACAGGAGACCTTCAAGATGGCCACACGAATAGTACAGC 9702
Qy 4503 ACTATATCATCGGATTCAGTCTACATGTCTATTTGTGCTGTGTGCTCTTGTCTA 4562
Db 9703 ATTATCTACATCGCATCTGTGTACACATCTTAAGCGGTGATCACTACCGTGGGGA 9762
Qy 4563 TCTCTGTAGGACATGATCATGACGCTTGATGCGCAAGAAAGAGATGCTCTGA 4622
Db 9763 TGAATATGGGCTAATCTTGTGACAGTGTATGTGCTGTAAAGCGCGGTGATGCTCTGA 9822
Qy 4623 CGGCATAAGCGCTTGGACCGAACGCAAGGATACCCACAGATTAAGCGGTTTTGTGCTGA 4682
Db 9823 CGGCATAAGCGCTTGGACCGCAAAACGCGTAAATCCCAATTCGCTGGCACTTTGTGCTCG 9882
Qy 4683 TTGCGCAACCAACGCTGAAACATTTGAGAACTTTGAACATCTGTGTGTTAAACAC 4742
Db 9883 TTAGGTGCGCAATGCTAAACGTTCAACGAGACATAGATTACTTTGTGTGCAACATC 9942
Qy 4743 AACGTTTCTCTGGGACAGTGTGATTCCTCTGGAGCGCTGTATTTCTGTTCCGCT 4802
Db 9943 AGCGGTTCTTCTGGGACAGTGTGATTAACCTTTTGGCGCTTTATCTGTTTAAATGCGCT 10002
Qy 4803 GCTTTTCAATGCTGATGCTTTTATTGTTGACAGCGCTCTGCTTGGGAAAGGTAGACG 4862
Db 10003 GCTGCTCTGCTGCTGCTTTTATTGTTGCTGGCGGCTTACTTGCGAAGGTAGACG 10062

Qy 4863 CTTTGCAGATGCGACACTGTGCAAAATGTTCCGGGATCCCGTAAAGCGGTGCTG 4922
Db 10063 CTTAGGAATGCGACACTGTGTTCCAAATGTGCAACAGATACCGTATTAAGCACTGTTG 10122
Qy 4923 AACGCGAGTTACGGGCACTTAACCTGGAATCAACGCTGTCTGATCGGAATTAAC 4982
Db 10123 AAAGGCGAGGTATGCCCCGCTCAATTTGAGATCACTGTCAATGCTCTCGGAGGTTTGC 10182
Qy 4983 CTTCACTAACAGAGATGATGATCTGTGCAAAATTTCCACACAGTCAATTCCTTCCACAC 5042
Db 10183 CTTTCAACCAACCAAGATGATTAATCTGCAAAATTCACACTGTGTGCTCTCCCAAAA 10242
Qy 5043 TTAATGCTCGGGTCCCTGAGTGCAGGATCTCTCAAGGCGGATTAACATGCGCGC 5102
Db 10243 TCAAATGCTCGGCTCTTGTGAATGTACGCGGCGCTTCAATGCAATTAACCTGCAAG 10302
Qy 5103 TTTTGGCGGTGTATCCCTTTCATGTGGGAGCGCAATGCTTCTGTACAGTGA 5162
Db 10303 TCTTGGAGGGGTCTACCCCTTTATGTGGGAGAGGCAATGTTTGTGGACAGTGA 10362
Qy 5163 ACACACACTGATGAGGCGTACGTGAGTTCCTCAAGCTGCACTATATGATCAAGCAG 5222
Db 10363 ACAGCCAGATGAGTGAAGCGCTAGCTGAACTGTCAAGATTCGCGTCTGACACGCGC 10422
Qy 5223 TCGCACTAAAGTTCAACAGCTGCTCTGAAAGTGGCGCGGTATGATTAAGGCAACA 5282
Db 10423 AGCGATTAAGTGTACACTGCGCGCATGAAATGAGATCTGATATGTATAGGGAACA 10482
Qy 5283 CACCGCGCACTGATACGTTTGTCAATGCGCTACAGCGGATTCCTCAAGGACCTGA 5342
Db 10483 CTACAGATTTCTAGATGTAGTGTAGTGTAACGAGTCAACACGAGAACGCTTAAAGACTTGA 10542
Qy 5343 AGGTATAGCAGGCGCATATACGCGCTTTTACACCTTTGACATTAAGTGTATCA 5402
Db 10543 AAGTATAGCTGACCAATTCATTCAGATCTGTTACGCGCATTCATTAAGTGTATTC 10602
Qy 5403 GAAAGGGCTTGTTTTACATACAGATTCCTCTGAGTATGAGGCTATGAACCAAGAGCGT 5462
Db 10603 ATGCGGCTGTGTGTACATATGACTTCCGGAATATGAGAGGATGAACCAAGAGCGT 10662
Qy 5463 TCGGCATATTAAGACATCTCGCTTGTATGCTACAGACATAGTACCCGCACTGACATAC 5522
Db 10663 TTGAGACATTTCAAGTACTCTCTTGTACAGAGATCTCATGCGCAGACAGACTTGA 10722
Qy 5523 GCGTCTGAAGCTTCTGTGAAGAACTACAGTCCCTTACACCCAGCAGTATCAGGCT 5582
Db 10723 GCGTACTCAAGCTTTCGCAAGAAACGTGATGTCCGTACAGCAGGCGCATCAGGAT 10782
Qy 5583 ATGAATGTGAAGAAACATCTCAGAGACGCCCTGCAAGAAAGCAACATTTGATGTA 5642
Db 10783 TTGAGATGTGAAGAAACATCTCAGGCGCGCCATCTGAGAAACCGCACTTTGGGCTGA 10842
Qy 5643 AAATTAAGTGAAGCTTCTGCAAGCGTCTAATCTGTGCTTACGCGGCAATCTTATCTGA 5702
Db 10843 AGATTCAGTAAATTCGCTCCGAGCGGTGAGCTGTATATAGGAAACATTTCCATTTCTA 10902
Qy 5703 TTGACATCCCTGATGACGCTTTTGTGATCATCAAGATTAACCAACAAATTTAGAACTTA 5762
Db 10903 TTGACATCCCGAAGCGCTTATATCAGACATCAGATGACCACTGCTCAACATGA 10962
Qy 5763 GCTGCAAGTGAAGACATGATTTATCTGACAGCTTGTGCTTCTAATCAATACAGT 5822
Db 10963 AATGTAAGTGAAGTGAATCTTATTAAGCAAGCTTGTGCGGAGATGCGACCTGAGT 11022
Qy 5823 ACAAGCTGACAGGAGAGCAATTTGTCAGTTCACCTCCACACAGCAGCTGTTTGA 5882
Db 11023 ATGTATCCAGCCGGAAGGTCAATGCCCCGTATATGATGATTCGAGACAGCAACTTCC 11082
Qy 5883 AGGAAGCAGACACATGTGATGCTGCGTGAAGACATTAACATTAATTTAGACATGCA 5942
Db 11083 AAGATGCAAGTACATGTCTCTGAGAAAGAGCGGTGACAGTACATTTAGACACCGCA 11142

QY 5943 GCCCAGCAAAATTTTATAGTTTGGCTATGCGGAGAGTGCACCTGCAATGCTGAAT 6002
DB 11143 GTCCACAGCGCAATTTATGATGCTGTGTGGAGAGAACATGATGCAATGAGAAAT 11202
QY 6003 GTAAACCAACCGGCGGACCAATTAATTGGAGAACCAATTAAGTCAGCAAGATTCCAG 6062
DB 11203 GTAAACCAACAGCTACCAATATCGTAGACACCCCGACAAAATATACCAAGATTTCAG 11262
QY 6063 CGGCACTTCCAAAACATCTTGAACCTGGCTGCTGCACTGTTTGGGGAGATCATCC 6122
DB 11263 CCGCATCTCAAAAACATCATATGAGTGGCTGTTGCTTTGCGCGCGCTGTGCG 11322
QY 6123 TCATTTGTATGAGCTTATAGTGTGTCTGAGCTATGCTTATTAACACACCTAGAT 6182
DB 11323 TATTATTTATAGACTTATGATTTTGTGCTGAGCATGATGCTGACACAGAAAGT 11382
QY 6183 GACTAGGCGGCGGACACTGA-CATAGCGGTAAACTGATGATCTTCCAGAGAGCTGTGT 6241
DB 11383 GACCGCTACGCGCCCAATGATCCGACAGCAAACTGATGATCTCCAGAGAACTGATGT 11442
QY 6242 GCATATGCGCAAGCGCGCT 6261
DB 11443 GCATATGATCATGAGCTGT 11462

RESULT 7
AAC64507
ID AAC64507 standard; DNA; 11703 BP.
XX AAC64507;
AC
XX 14-FEB-2001 (first entry)
DT
XX
DE Alphavirus SinchironLP virus nucleotide sequence.
XX
XX Alphavirus; Sindchiron virus; SinchironLP virus; immune response;
KW infection; human dendritic cell; immunostimulatory; cytostatic;
KW virulence; fungicide; antibacterial; antiparasitic; vaccine; cancer;
KW pathogen; antigen presenting cell; ds.
XX
OS Alphavirus.
XX
XX MO200061772-A2.
PN
XX 19-OCT-2000.
PD
XX 14-APR-2000; 2000MO-US10722.
PF
XX 14-APR-1999; 99US-0129498.
PR 09-AUG-1999; 99US-0148086.
PR 22-MAR-2000; 2000US-0191363.
XX
XX (CHIR) CHIRON CORP.
PA
PI Pojo JW, Dubensky TW, Frolov I, Gardner JP, Otten G, Barnett S;
PI Driver DA;
XX
XX WPI; 2000-619231/59.
DR
XX
PT New alphavirus that infects human dendritic cells for use in generating
PT an immune response to pathogenic agents such as bacteria, viruses,
PT fungi, parasites and cancer and for biological assays -
XX
XX Claim 10; Fig 2C; 83pp; English.
XX
XX The present invention describes an isolated alphavirus (AV) which
CC infects human dendritic cells and is not of American Type Culture
CC Collection (ATCC) number VR-2526. AAC64506 and AAC64507 represent the
CC nucleotide sequence of the specifically claimed Sindchiron virus and
CC SinchironLP virus. The new AVs have immunostimulatory, cytostatic,
CC virulence, fungicide, antibacterial and antiparasitic activities and
CC can be used in vaccines. The AVs are used to infect dendritic cells,
CC preferably human cells. A heterologous sequence can be introduced and

CC expressed in human macrophages or antigen presenting cells in vivo and
CC in vitro, for use in biological assays. The AV-based vector systems
CC are used to generate an immune response to cancer or a pathogenic
CC agent, such as, bacteria, fungi, parasites or viruses. The AV can be
CC used to infect human dendritic cells; macrophages or antigen
CC presenting cells that previously could not be infected using an AV or
CC AV variant. The AV vectors are targeted directly to antigen presenting
CC cells.
SQ Sequence 11703 BP; 3308 A; 3050 C; 2911 G; 2434 T; 0 other;
XX
XX
XX Query Match 36.7%; Score 2382; DB 21; Length 11703;
XX Best Local Similarity 64.6%; Pred. No. 0;
XX Matches 3694; Conservative 0; Mismatches 1985; Indels 41; Gaps 8;
QY 580 ACTGACGATGAAAGCGGAGCGGATATATTTCTCATCGGAAACAGGCCAGTCACTTC 639
DB 5746 ACTGACTAACCGGGGTAGTGGTATATATTTTGCACGACAGGCCCTGGCCTTGC 5805
QY 640 AACGAAATCAGTACGTCAATGTAACTACAAAGAACCTATATTTGGATGGGCGCTCAG 699
DB 5806 AAAAGAGTCCGCTTCTGAGAACCACTTACAGAACCACTTGGAGCGCATGTCTCG 5865
QY 700 AGAAGTATTAAGCCCGCGCGCTCGATCTGAAAGAGAAATGTATACAGAAATCTGC 759
DB 5866 AAAAGATTCATGCCCGGTGCTCGACAGCTGAAAGAGAACCACTCAAACTCAGGTAC 5925
QY 760 AATTATGCGCTCTGAGAAAGAAATAGAGACAGTATCACTACGAAAGTAAATATGA 819
DB 5926 AGATGATGCCCAAGCAAGCCAAAGATAGTACCACTCTGTAAGTAAATATCA 5985
QY 820 AAGCAATTACAGCGGAGGAGCTCACTTCTGATTTGGGACATATCTATCAGAAAGA 879
DB 5986 AAGCATACCACTGAGGAGTACTGTGAGACTACCACTGTATTAATCTTCCACAG--- 6042
QY 880 ATCTGTGAGTGTATACAGATCAATTATCTGTACCAATCTACTCTGCAACGCTAATTA 939
DB 6043 ATCAGCCAGATGCTATATGATATCACTATCGAAACATGTAATCTTCCAGTACCG 6102
QY 940 ACAGGTTTACATCTGACAGAGTCCGGTTAAAGCTGCACTTACTTATCCAGAGATT 999
DB 6103 CGAACTACTCCGATCCACAGTTCCTGTAGCTGTGTAAACATATCTGATGAGAACT 6162
QY 1000 ACCCTACAGTACCAAGTATTTATACAGATGATAGATGCGATCTGACATGAGG 1059
DB 6163 ATCCGACATGACATCTTATCAGATTTACTGACGATGAGATCTTACTGATATGAG 6222
QY 1060 ACGGCGATCGTGTCTGATACAGCCACTTTTTCGCGCTAAACTGAGAACTTACC 1119
DB 6223 ACGGACAGTGGCTGCTGATGATGACACTTCTGCGCTAGAGTTAGAGTTACC 6282
QY 1120 CAAGAGCATAGCTATTTTTCGAGCCAGATTAAGATACAGCCGCTCCATGCTTATACA 1179
DB 6283 CGAAAAAATAGATATAGAGCCCGCAATATCCGAGTGGCTTCAATCAGATGAGA 6342
QY 1180 ATACATTACAAATGTATGCTGCTGCACTATTAAGAAATTGCAACCTTACCAAAAGC 1239
DB 6343 ACAGCTACAAATGTGCTCATGCTCCGCACTTAAAGAAATGCACTTACGCAATGAC 6402
QY 1240 GAGATTAACCTGTCTTATGATTCGCGGATTTAATGTTGATTTTCAAGAAATACGAT 1299
DB 6403 GTGAAGTCCCAACTGAGTACGACGACATTCATATGCTGATGCTTTGAAATATGAT 6462
QY 1300 GCAATGATAGTACTGGGATACCTTTTGGCGATTAACCTATTCGCTTAACTACAGAAAG 1359
DB 6463 GTAAATGAGATATTTGGAGAGATTCGCTCGAAGCCAAATATGAGTTACCTAGTTTG 6522
QY 1360 TTAGCAATATGATGACAAAGTGAAGGCGGAAAGCAAGCACTGTTTGGAAATCTC 1419
DB 6523 TCACCGCATATGATGATGATGAAAGGCTTAAAGGCGCGCACTATTTTCAAGAGAGT 6582
QY 1420 ATATATCTAAACCGTTGAGAGATACCAATGATGATCAATTCATGATGATTAAGAGAG 1479

Dh 6583 ATAATTGTCCTCCATGTCAGAAAGTGCCTATGATGATTCGTATGATGACATGAAAGGG 6642
Qy 1480 ATGTCAAAGTTACTCCGGGACGAAACATACAGAGAGCGGCTTAAGTGTGATATTC 1539
Db 6643 ACGGAAAGTTACACGAGGACGAAACACAGAAAGAACCCAAAGTAAAGTGTATAC 6702
Qy 1540 AGGCTGAGATCCCTGCTACCGCTTACCTTGGGGATCCATCGGGAAATTAAGTCGTA 1599
Db 6703 AAGCCGAGAACCCCTGGCGACTGCTTACTTATGCGGATTCACCGGAAATTAAGTCGTA 6762
Qy 1600 GACTGAATGCGGTCCTTTCGCAAAATATCATCTCTTTCGATGTCAGCGGAGATT 1659
Db 6763 GGCCTTACGGCCGCTGCTTCCAAACATTCACGCTTTTTCATGTCGCGGAGATT 6822
Qy 1660 TTGATGCGATTATGCTGAAATTTCCACGCGGACCCGATTTGGAAAACGACATCG 1719
Db 6823 TTGATGCAATCATGACGAAACACTTCAGCAAGGCGACCCGCTACTGGAGCGGATTCG 6882
Qy 1720 CGTCGTTGATTAAGGCAAGACGACGCTATCGCATTTGCGCGTTGATGATCCCTGAGG 1779
Db 6883 CATCATTCGACAAAAGCCAGACGACGCTATGCGGTTAACCGGTCTGATGATCTTGGAGG 6942
Qy 1780 ACTTAGTGTGACCAACCGCTCTTGAATTTGATAGAGCGCGCTTCGCAATATCAT 1839
Db 6943 ACCTGGGTGTGATCAACCACTACTCGACTTGATCGAGTGCCTTTGGAAATATCAT 7002
Qy 1840 CTGTCACCTACCTACAGGAACGAGTTAAATTTGGTCATGATGAAATCCGGTATGT 1899
Db 7003 CCACCCATCTACCAAGGCTACTGCTTAAATTCGGGGCGATGATGAATCCGGAATGT 7062
Qy 1900 TCTTAAAGCTGTTGTCAACACACTAGTCATATCATGATGCTAGCAGATACCTACGTC 1959
Db 7063 TCTTCACACTTTTGTCAACACACTTTGAAATGCTGATTCGCAAGCAGAGTACTAGAG 7122
Qy 1960 AAGCGTTAACCAAGTCAGCGCTGCGCGCTCTATCGCGAGATACATAGTGCATGTC 2019
Db 7123 AGCGGCTTAAACGCTCCAGATGTGACGCGTTCATTTGGCGACGACATCATCATGAG 7182
Qy 2020 TCGTCTCCGACACCTTGATGCGGAGAGATGCGGCATCTTGGCTGAACATGGAATMAAA 2079
Db 7183 TAGATATCGACAAAGAAATGCTGAGAGTGCACCTGCTCAACATGGAAGTTAGA 7242
Qy 2080 TTATGATGACGATATTGATTAAGACACCTTCTGCGGGGATTTTCCGTGTCG 2139
Db 7243 TCAATGACGAGCATCGGTGAGAGACCACTTCTGCGCGGATTTTCTTTCGACG 7302
Qy 2140 ACCAGATTAACAGGACAGCTGACAGATGCGACACCTCTTAAAAAGGCTTTTAAAGCTTG 2199
Db 7303 ATTCGGTTACTTCCACAGCGTGCCTGCGGACCCCTGAAAAAGGCTGTTTAAAGTTGG 7362
Qy 2200 GAAAAACATTGGCAGTGTGATATACCCAAAGCTGCGACCGCGCGGCACTGCATGATG 2259
Db 7363 GTAAACCGCTCCAGCGGACGACGAGCAAGACAAAGACGAGACGCGCTCTGTTAGATG 7422
Qy 2260 AAGCAATGCGATGAAACAGATTGGAATTAAGGACGATTAAGGACCGCTAGAAATCA 2319
Db 7423 AAACAAAGGCGTGGTTTAAAGTAAAGTAAACAGGCACTTTAGCAGTGCCTGACGACCC 7482
Qy 2320 GATAAGATCATATCTGGCAGGCTGATCATCAGCTCTCTGTCCAAGCTTTACCGGAAACG 2379
Db 7483 GGTATGAGGTAGACAAATTTACCTGTCTACTGTGACATGAGAACTTTTCCGAGAGCA 7542
Qy 2380 TTAAAGACTTCAAGAGCATTAAGAGGAGCCCAATCACTCTTAAGGCTGACCTTAATAG 2439
Db 7543 AAAAGACATTCAGGACCATGAGAGGGGAAATTAAGCATCTCTACGGTGTCTTAAATAGT 7602
Qy 2440 TGAGGTATAGACAGCAGCTAC-----CCACGGGAGAAATGTTTCCATACC 2486
Db 7603 CAGCATATGATTTTCACTCTGACTAATTAATAACACACCAACCACTGAATATGAGATTTCT 7662
Qy 2487 CTCAGCTGAATTTTCAACAGATTAACTTAACTTCCGATGCTTAAACCGAGATCAACC 2546
Db 7663 TTAAATGCTGCGCGCGCGCCCTTCCGCGCCCACTGCATGTGAGAGCGCGGAGAA 7722

Qy 2547 CTCTAGAGGCGCGTGGAGGCGGTTTCGAGCCCGCGCTGCTCAATGAGATCTTA 2606
Db 7723 GAGAGCAGGCGGCCGATGCTGCTCCGACAGCGGCTGAGCTTCCAAATCAGCACTGA 7782
Qy 2607 GAGGTGATATGATCACTTGAATTTCAACCAAGATCACTTAATCCGCGCAGTCCA- 2665
Db 7783 CCAAGCCGTCAATGCGCTAGTATGATGAGCAGGCACTAAGCTCAACCCCGACGTCCAC 7842
Qy 2666 -----CCGCAAAAGAAAGAGAGTGTCTTAAAGCCAAACCTTACGCTTAAAAAG 2720
Db 7843 GCCCGCACCGCGCGAGAAAGCAGCGCGCCCAAGCAACCAAGGAGCCAGAAACCA 7902
Qy 2721 AGAAGCAGAGCAAGCAAGAGAGCAAAAC--GCAAGCTTAAACAGGAGAAACGACGTA 2777
Db 7903 AAAGCAGAGAAAGAAAGAAAGCAACCTTCAAAACCCCAAGGAAAGACAGCGCA 7962
Qy 2778 TGTGTATGAAGTTGAGTCCGACAAACATTTCCGATCATG---CTGACGCGCAAGTGA 2834
Db 7963 TGGACTTAAGTTGAGAGCCGACAGATCGTTGACGTCAAGAACGAGAGATGTCA 8022
Qy 2835 ATGATATGCTCTGCTGTGTGAGGAAAGCTGATGAACCACTCCAGCTTGAAGAAAA 2894
Db 8023 TCGGACAGCACTGCGCATGAGAAAGGTAATGAACCTCTGCACTGTAAGAAAGCA 8082
Qy 2895 TTGATATGAGCAATTAGCGCGCGTGAATTTGAAGAGCTAGCATGTACGACTGGAGT 2954
Db 8083 TCGACACCTGTGCTTCAAACTCAAAATTTCAAGTGTGACATACGATGAGAT 8142
Qy 2955 ACGGCGAGCTTCCCGAAGACATGAATACAGACGCTGCAGTACACCGAGCAACAC 3014
Db 8143 TCGACAGCTTGGCCAGTCAATGAGAGTGAAGCATTCACCTACACAGTGAACACCCG 8202
Qy 3015 CGGCTTCTTCAACTGGCACCAAGCGGAGTCCAGATGAGAAATGGAGATTTACCTGAC 3074
Db 8203 AAGGATTTCTAATCTGGCACACGAGCGGTGAGATGATGAGAGTTTAAACATTC 8262
Qy 3075 CGAGAGAGTGGCGGGGAAAGGACAGCGGAAAGACCGGATCTGGAACAAGAGCGAG 3134
Db 8263 CTGCGAGATGAGAGGACAGAGAGACAGCGGTGCTCCGATCATGATATCTCCGTCGG 8322
Qy 3135 TTGTGCTATTTGTTAGAGGTGCAAAATGAGGACGCGCTACGCGCTTTCAGTGTCA 3194
Db 8323 TTGTGCGATAGTCTCGTGGAGCTGATGAAGAACAGAACGACCTTTCGCTGCTCA 8382
Qy 3195 CTGGAACCAAGAAAGGCTGACATTTAGGATACCCCGAAGTTTGAACCTGTGT--- 3251
Db 8383 CTTGGAATTAAGGAGAAACATTAAGACGACCCCGGAAAGGACAGAAAGTGTCCG 8442
Qy 3252 -----CACTAGTTACAGCGCTATGCGTCTTTCGATGTCAGTCCAGTCCGACAAAC 3305
Db 8443 CAGCACCTAGTCAAGCGCAATGTGTTGCTCGAAATGTAGCTTCCATGCGACCGCC 8502
Qy 3306 CACCGTGTCTATTCACTGACCGCAAGACCACTGACGCTGCTGAAGAGACGTCG 3365
Db 8503 CGCCCATGCTATTAACCCCGGAACCTTCCAGAGCCCTCGACATCTTGAAGAGAACGTA 8562
Qy 3366 ACAATCAAAATTACGACCGCTGTGAGAACGCTTTGAAATGTCAATC---GCGGCG 3422
Db 8563 ACCATGAGGCTTCAATACCTGCTCAATGCAATTTGCGGTGCGGATCTCTGCGAGAA 8622
Qy 3423 CCAAGCAGATTAACGATGATCTTCACTGACAGCTGCTTACCTGCGGTTTTCGCGCT 3482
Db 8623 GCABAAAGCGTCACTGAGACCTTTTACCTGACAGCCCTTACTTGGGACATGCTCTGT 8682
Qy 3483 ATTGCAACACTCAACGCGCTGTTTTCAGCCCAATTAATAATGGAACGTTGGAGCAAT 3542
Db 8683 ACTGCCACATCATGAAACGCTGCTTACGCTGTAAAGATGAGAGGTCTGGAGACGAAG 8742
Qy 3543 CTGATGATGATGATTAAGATCAAGGTCTGGGCACAATTCGCTACATCAGCAGGCA 3602
Db 8743 CGGACGATTAACCAATACGATACAGACTTCCCGCGAGTTTGGATTCGACCAAGCGGAG 8802

QY 3603 CTGGGATGTACCAAAATTCGGTACATGTCTTTCGACCAAGCATGACATCAAGAG 3662
Db 8803 CAGCAAGGCGAAACAGTACCGCTACATGTCCCTTAAGCAGAGATCAACCGTTAAAGAG 8862
QY 3663 ACAGATGAGAAATAGTATCAGCAATCTGACCCCTGCGTGTCTTGGCCCAAG 3722
Db 8863 GCACCATGATGACATCAAGATTACACCTCAGGACCGGTGTGAAGGCTTACTCAAG 8922
QY 3723 GGTACTCTGTCTTACTCAATGTCCCTCCAGGTGACAGTGTACCGTATCAAGAG 3782
Db 8923 GATCTTCTCTCCGCAAAATTCCTCCAGGGACAGCGTAAAGGTGACATAGTAGTA 8982
QY 3783 GAGCATCTGAGAAATTCATGACCCGTGAGAAAAGATCAGAGAAATTTCTCGTAGAG 3842
Db 8983 GCAACTGACGAACGTCATGTACCTGGCCGCAAGATTAACCAAAATTCGTGGACCGG 9042
QY 3843 AGGAGTACTTGTCCCAACCCGTCATGGAAGCTGTAAAGTCCACGTTACATCACT 3902
Db 9043 AAAATATGATCTACCTCCCGTTCAAGGTAAAAAATTCCTTGACAGGTGTGACCGTC 9102
QY 3903 TGAAGAGACGTCCTCCGGGTACATACCATGACAGGCGACCGCATATAGT 3962
Db 9103 TGAAGAGAACATGACAGCTACATCATGACAGGCGGAGCCGACGCTTATACAT 9162
QY 3963 CCTATCTGAGGAGCGTCAGGCGAAGTGTACATTTAAACACCTTCTGGCAGAAAGTCA 4022
Db 9163 CCTACTGGAAGAAATCATAGGGAAGTTTACGCAAAAGCCGCACTGGGAAGAACATTA 9222
QY 4023 CCTACGAATGTAGTGTGCGCATACAGACAGTATGCTGAGCACGCGAACGAAGATGA 4082
Db 9223 CGTATGAGTGCAGAGCGCGGCTCAAGACCAAGACGTTTGCACCGGCAAGAAATCA 9282
QY 4083 ACGGTGACATTAACCAAAACGTGATCTTCTTAAGAGCGACCAAGGAAATGGGTCT 4142
Db 9283 CTGGTGTGACCGGCATCAAGAGCTGCTGCTTAAGAGCGACCAAGGAAATGGGTCT 9342
QY 4143 TCAACTCGCCGATCTTATTAAGGACACAGACACCTCAGTGCAGATTAATTTGACATTC 4202
Db 9343 TCAACTCAACCGGACTTGTATGACATGACAGACACACCGCCCAAGGAAATTTGCTTGC 9402
QY 4203 CATTCGCTTGAACCGACAGTCTGCGGTTCGCTTACGTCAACGCTTACAGTACGA 4262
Db 9403 CTTCACATTTATCCCGAGTACTGATGTCCTGTTGCCACCGCGCAATGTAATAC 9462
QY 4263 AGTGTGAAGGACATCAACCTCCACCTGACCTGACATGACCAACATTTGCTGACAAAGA 4322
Db 9463 ATGGCTTTAAACATCACTGAGCTTCCAAATTAGATAGACCACTTGACATTTGCTCACACCA 9522
QY 4323 GAAAAATGGGCTGCGAGACAGACGACAGAGATGATTAAGGGTCTACATCCAGGA 4382
Db 9523 GGAAGCTAGGGGCAAAACCGGAAACCACTGAATGGAATGCTCGAAAAGAGCGTACGAA 9582
QY 4383 ATTTTCTGTGGGGGAGAAAGGCTGAGTACGTATGGGGTAAACCATGAACCATGACAG 4442
Db 9583 ACTTCAACGTCGACCGAGATGGCTGGAATACATATGGGGAATCATGAGCCAGTGAAGG 9642
QY 4443 TCTGGGCCCCAGAGTCTGCGACACAGGCAACCATGATGCTCGCATGATCATATCC 4502
Db 9643 TCTATGCCCAAGATCAGACACAGGACCCCTCAAGATGGCCACACCAAAATAGTACAGC 9702
QY 4503 ACTATATCATGCGCATCAAGTCTACATGTCATGTGTGTGTGTGTGTGTGTGTGTGT 4562
Db 9703 ATTACTACATGCGCATCTGTGTACATCATTTTGTGCGGTGCGATCAGTACCTGTGGGA 9762
QY 4563 TCTGTGAGGACATGCAATCAAGAGCTTGCATGCGCAAAAGCAAGAAAGAGACTGCTGA 4622
Db 9763 TGATGATGGCGTAACTGTGCAAGTATGTCCTGTAAGCGCGCGGTGAGTGTCTGA 9822
QY 4623 CGCCATAGCGCTTGCACCGAAACGAGGATCCACAGCATTTAGCGGTTTGTGTCTGA 4682
Db 9823 CGCCATAGCGCTTGCACCGAAACGCGGTAATCCCAACTTGTGTGTGTGTGTGTGTGT 9882
QY 4683 TTGCGCAACCAACGCTGAACATTTGGAGAAATTTGAACCATGTGTGTGTGTGTGTGT 4742

Db 9883 TTAGTGGCGCAATCTGAAACGTTACACGAGAACATGATGATCTGTGTGGAACAGTC 9942
QY 4743 AACGTTTCTGTGGGACAGATTGTGATCTCTGTGGACGCTTGTATTTCTGTCGCT 4802
Db 9943 AGCCGTTTCTGTGGGTCAAGTGTGATACCTTTGTGGCGCTTTCATGCTTAAATGCGCT 10002
QY 4803 GCTTTCAAGTGCATGCTTTTATATGTTGTGACAGGCGTCTGCGTGGGGAAGTAGAG 4862
Db 10003 GTTGTCTGTGCTGCTGCTTTTATGTTGTGTTGCGGCGCTTACCTGGGGAAGTAGAG 10062
QY 4863 CTTTGAACATGCGACACACTGTGCCAAATGTTCCGCGGATCCCTGTAAAGCGTGTG 4922
Db 10063 CCTAGGAACATGCGACACACTGTGCCAAATGTTCCAGATACCTGTAAAGCACTTGTG 10122
QY 4923 AACGCGAGTTTACCGCGCATTAACCTGAGATCAAGGTGTCTCATGGAATTAACAC 4982
Db 10123 AAAGGCGAGGATATCCCGCTCAATTTGAGATCATGTCTCATGTCTGAGGTTTGC 10182
QY 4983 CTTCAATTAAGAGATGACGTGACCTGCAATTCACACAGTCAATTCCTTACACAGAG 5042
Db 10183 CTTCCACCAACCAAGATACATTAATCTGCAAAATTAACAATGTGTCTCCCTCCAAAAA 10242
QY 5043 TTAATGTGCGGAGTCCCTGAGTCAAGGACCTTCAAGCGGATTAACATGCTCCG 5102
Db 10243 TCAATGTGCGGCTCTTGGAAATGTCAAGCGCGCTTCAATGACATTAATCTGCAAG 10302
QY 5103 TTTTGTGCGGTGTGATCCCTTTCATGTGAGGAGCGCAATGCTTCTGTGACAGTGA 5162
Db 10303 TCTTGGAGGAGGTCTAACCTCTTATGTGGGAGAGGCCAATGTTTGGCAGATGAGA 10362
QY 5163 ACACACATGATGAGGAGGATACGTGAGTTCGTCTCAGACTGCACTATAGTACAGAG 5222
Db 10363 AACGCAATGATGAGGAGGATACGTGAGTTCGTCTCAGACTGCACTATAGTACAGAG 10422
QY 5223 TCGCATTAAGATTCACACAGCTGCTGAAAGTGGCGTGTATGATATCGGCAACA 5282
Db 10423 AGGCGATTAAGTGTGACATGCGCGAGTGAAGATGAGCTGTGTATGTATGAGGGAACA 10482
QY 5283 CCACCGCGACCTGGATACGTTTGTCAATGGCGTCAAGCCAGTTCCTACGGGACTGA 5342
Db 10483 CTACCAATTTCTGTATGTGATGACGTAAGAGGATGACACACAGGAACGCTTAAGACTGA 10542
QY 5343 AGTCTATGACGCGCGGATATCAAGCGCTTTTCAACCTTGAACATAGGTGTGCATCA 5402
Db 10543 AAGTATGCTGGAACAAATTTCAAGCATGTTTACCGCATTCGATATTAAGGTGTATTC 10602
QY 5403 GAAAGGCGCTGTTTAACTACAGCTTCCCTGAGTATGAGCTATGAACAGAGGCT 5462
Db 10603 ATCGGCGCTGTGTATCAACTATGACTTCCCGGAATATGAGCAATGAACAGAGGCT 10662
QY 5463 TCGGCAATTAACGATCTCTGCTGTATGCTACAGACATAGTACCGGACATGACATAC 5522
Db 10663 TTGAGACATTAACGATCTCTGCTGTATGCTACAGAGATCTATGCTGACAGACATTA 10722
QY 5523 GCGTCTGAAGCTTCTGTCAAGAACATCAAGTCTCCCTTACACCCCAAGCATATAGGT 5582
Db 10723 GCGTACTGAAGCTTCTGTCAAGAACATCAAGTCTCCCTTACACCCCAAGCATATAGGT 10782
QY 5583 ATGAATGTGGAAGAACATCAAGACAGACCCCTGCAAGAAACAGACATTTGATGATGA 5642
Db 10783 TTGAGATGTGGAAGAACATCAAGACAGACCCCTGCAAGAAACAGACATTTGATGATGA 10842
QY 5643 AAATGAAGTGAAGCTTGTGAGAGGCTTAACTGTGCTTACGAGGACATCTCTATCTGA 5702
Db 10843 AGATTGAGTGAAGCTTGTGAGAGGCTTAACTGTGCTTACGAGGACATCTCTATCTGA 10902
QY 5703 TTGACATCTGATGAGCACTTTTGTGAGATATCAAGATTAACCAAAATTTGAAGTTA 5762
Db 10903 TTGACATCTGATGAGCACTTTTGTGAGATATCAAGATTAACCAAAATTTGAAGTTA 10962
QY 5763 GCTGACAGTGAAGCACTTATTTATCTGAGACTTTGTGAGTCTCTTAACTATTAAGT 5822

Db 10963 AATGTGAAGTCAGTGAAGTCACTTATTCAGCAGACTTCGGGGAGTGCCACCTGCAGT 11022
Qy 5823 ACAAGGTGACAGGAGGAGCATTTGTCCAGTTCCTCCATCCACGACAGCTTTTGA 5882
Db 11023 ATGTATCCGACCGGAGGATGATGCCCCGTACATTCGATTCAGACAGCAACTCTCC 11082
Qy 5883 AGGAAGGACACACATGTGACTGCCGTAGGACATTAACACTATTTTAGACATCGA 5942
Db 11083 AAGAGTGCAGATGATCTCTGAGAAAGAGGGGTGACATCTTTAGCACCAGCA 11142
Qy 5943 GCCCACAAGCAATTTTATGATTGCTGATGCGCAAGAGTCCACTGCAATCTGAT 6002
Db 11143 GTCCACAGGCGAATCTTATCTGATGCTGTGTGGAAAGACAATGCAATGCAGAT 11202
Qy 6003 GTAAACACGCGGCGACCACTAATTGGAAACCAATTAAGTGCACCAATTCAGG 6062
Db 11203 GTAAACACACGCTGACCATATTCGTAGACACCCGCAAAAATGACCAAGAAATTCAG 11262
Qy 6063 CGGCAATTTCCAAAACATCTTGAAGTGGCTGTGCACTGTTGGGGAGCATATCC 6122
Db 11263 CCGCATCTCAAAAACATCATGAGTGGCTGTGCTTTGGCGGCGCTCTGTCG 11322
Qy 6123 TCATGTTGTGAGCACTTATGATGTTGTCTGACCTCTAGCTTATAAACACAGTATGAT 6182
Db 11323 TATTAAATTAATGAGACTTATGATTTTGTCTGACATGATGCTGACTAGCAGAGAT 11382
Qy 6183 GACTGACGCGGAGACATGCA-CATAGCGGTAAACTGATCTTCCGAGAGAACTGCT 6241
Db 11383 GACCGCTACGCGCCCAATGATCGACGCAAAACCTGATGATCTTCGAGAACTGATGT 11442
Qy 6242 GCATATGCGCACGCGCGCT 6261
Db 11443 GCATATGCGCATCAGGCTGCT 11462

RESULT 8
AAT47668
ID AAT47668 standard; cDNA; 11663 BP.
XX
AC AAT47668;
XX
DT 19-MAY-1997 (first entry)
XX
DE South African Arbovirus 86 RNA transcript cDNA.
XX
KM S.A.A86; Sindbis virus; attenuation; vaccine; diagnosis; antibody;
KM vector; Ockelbo; polyarthrititis; ss.
XX
OS South African Arbovirus No. 86.
XX
FH Key location/Qualifiers
FT mutation 215
FT /*tag= a
FT /note= "a guanine for adenine substn. at nucleotide
FT 215 is a preferred silent mutation used in
FT recombinant DNA constructs of the
FT invention"
FT
FT mutation 3863
FT /*tag= b
FT /note= "a guanine for cytosine substn. at
FT nucleotide 3863 is a preferred silent
FT mutation used in recombinant DNA constructs
FT of the invention"
FT
FT mutation 5984
FT /*tag= c
FT /note= "a guanine for adenine substn. at nucleotide
FT 5984 is a preferred silent mutation used in
FT recombinant DNA constructs of the
FT invention"
FT
FT mutation 9113
FT /*tag= d
FT /note= "a cytosine for thymine substn. at
FT nucleotide 9113 is a preferred silent

FT mutation used in recombinant DNA constructs
FT of the invention"
XX
XX MO9637220-A1.
PN
XX
PD 28-NOV-1996.
XX
PD 22-MAY-1996; 96MO-US07457.
XX
PD 23-MAY-1995; 95US-0446932.
XX
PR (UNCL-) UNIV NORTH CAROLINA.
PI Davis NL, Johnston RE, Simpson DA;
XX WPI: 1997-020937/02.
XX
XX New recombinant DNA encoding infectious South African Arbovirus 86
PT RNA transcript and related mutants - under control of heterologous
PT promoter; useful in prodn. of attenuated vaccines and to generate
PT diagnostic antibodies
XX
PS Claim 2; Page 13-19; 27pp; English.
XX
XX A cDNA clone (AAT47668) codes for an infectious South African
CC Arbovirus No. 86 (S.A.A86) virus RNA transcript. Novel
CC recombinant DNA constructs comprise the cDNA, pref. mutated to
CC include attenuating and silent mutations, and an upstream
CC heterologous promoter (pref. T3, T7 or Sp6). Infectious
CC attenuated viral particles produced from cells transfected with the
CC RNA transcripts encoded by such cDNA clones are useful in live
CC attenuated vaccines, e.g. against diseases caused by related
CC viruses such as Ockelbo, and can also be used as vectors to express
CC antigens of other viruses. The cDNA clone can be administered to
CC animals to raise antibodies for diagnosis of S.A.A86 infection.
XX
SQ Sequence 11663 BP; 3297 A; 3104 C; 2903 G; 2359 T; 0 other;
Query Match 36.7%; Score 2381.6; DB 18; Length 11663;
Best Local Similarity 64.4%; Pred. No. 0;
Matches 3742; Conservative 0; Mismatches 1994; Indels 74; Gaps 9;
Qy 580 ACTGACCGTATGAAAGCGGAGCGGTATATTTCTCATGGAACAGGCGCAAGTCACTTC 639
Db 5707 ACTGCTAACCGGGGTAGTGGGTACATATTTTGAAGCAGACAGGCGCTGGCATTTGC 5766
Qy 640 AACGAAATACATGATGCTCAATGTAATCAACAACTATATTGGATCGGCGCTCATG 699
Db 5767 AAAAGAAAGTCCGTTCTCAGAACCCAGCTTACAGAACCCAGCTTGGACGCAATGTTCTGG 5826
Qy 700 AGAAGTATTAAGCGCGCGCGCTGATCTCGAAAGAGAAATGTTACAGAAAGTGC 759
Db 5827 AAAGAAATCTACGCGCGCGGTGTCACACGTGGAAGAGAACCTCAAACTCAGGTACC 5886
Qy 760 AATTATCGCGCTCTGAAAGAAATGAAAGCAGTATCAATCAAGAAAGTGAATAATGA 819
Db 5887 AGATGATGCCACCGGAGAACCAAGAAAGCAGTACAGTCTCGAAAGTGAACCAAGA 5946
Qy 820 AAGCAATTAACGCGGAGCGACTATTTCTGGATTGGCACAATATCTATCAGAAATGA 879
Db 5947 AAGCATATACCACTGACGCGACTGTTTCAAGGCTACACATGTAATACCTGCCACAG--- 6003
Qy 880 ATCTGTGCGAGTGTAAAGAGTCAATATCTCTGACCAATCTACTGCAACGTAATTA 939
Db 6004 ATCAGCGAATGTATTAAGATCACTTACCGAAACATGCTATTCAGAGGTACACG 6063
Qy 940 ACAGTTTACATCTGACAGAGTCCGGTTAAACGTCAACTTATGTTATCCAGAGATT 999
Db 6064 CGAATCTACTGACCCAAAGTTTGTAGCTGTTTGAACAATATCTGATGAGATT 6123
Qy 1000 ACCCTACAGTACCGATTTATGTAACAGATGAATACATGATGCTATCTTGACATGCTGG 1059
Db 6124 ACCGACGATGATCTTATTCAGATCACCGACGATGATGATGCTTACTTGATATGTTAG 6183

OY	1060	ACGGGCACTCGTGGCTGTCTAGATTCAGGCACTTTTGTCCGGCTTAACTGGAAGCTTACC	1113
Db	6184	ACGGGACAGTGGCTTGCTTAGATCTGMACTTTTGTCCCGCAAGTTGAAGTTTACC	6243
OY	1120	CAAAAGACATAGCTATTTTGGACGCCAGAGATTAAGATCAGCCGCTCCATCGCTTATACAGA	1179
Db	6244	CGAAAAGACACGAGTATAGAGCCCAACATCCGAGTGGGTTTCATCAGCATGTGACA	6303
OY	1180	ATACCTTACAAATATATTGGCTGACGATCTTAAAGGAATTGCACGTTACCCAAATGC	1239
Db	6304	ACAGCTTGCAAAACCTGTCTCATTTGCCGGACTTAAAGAACTGCACCTCCACACAAATATGC	6363
OY	1240	GAGAAATTACTGTCTTAATATTCGGGGCGATTAAATGTGATTGTGTTCAAGAAATACGAT	1299
Db	6364	GTGAAGTCCCAACACTGGACCTCAGGAGATTTCAACGTTGAATGCTTTGAAATATATGAT	6423
OY	1300	GCAATGATAGACTACTGGGATACCTTTGGCGATAACCTTATTCGGCTTAACTACAGAAAC	1359
Db	6424	GCAATGAGAGATATTGGGAGAGATTTCCCGAAAGCCAAATTAGATCTACTAGATTGG	6483
OY	1360	TTAGCGAATATGTGACAAAGCTGAAAGGGCCGGAAGACAGCAGCATGTGTTCCGAATCTC	1419
Db	6484	TTACCGCATAGCTGGCCAGACTGAAAGGCCCTTAAGGCCCGCAGCACTGTTCCAAAGAGCC	6543
OY	1420	ATTAATCTTAAACCGTTTGACAGAGATACCAATGATCTCAATTTCGTATGATGATTAAGAG	1479
Db	6544	ATTAATTTGCTCCCATTCGCAAGAAAGTGCTTATGATATAGTTGATCATGACATGAAGAG	6603
OY	1480	ATGTCAAAAGTTACTCTCCGGCACGAACATPACAGAGAGCGGCTTAAGTGCAAGTTATTC	1539
Db	6604	ACGTGAAGATTACACTGTGGCACGAACACACAGAAAGAAAGCCAAAGTACMACTGATATAC	6663
OY	1540	AGGCTGCAGATCCCGCTTGCTACCCGTTACTTTGGGGATCCATCGGGAAATTAGTCCGTA	1599
Db	6664	AAAGCCGACGAACCCCTGGCGAGCCGTTTACTATGGGGATTCACCGGGAGTTAAGTGCGCA	6723
OY	1600	GACTGAATGCGGTGCTTGTGCCAATATCACTCTTTCGACATGTCAAGCGGAAGATT	1659
Db	6724	GGCTTACAGCGGTTTGTCTTACCACCAATTCACAGCTCTTGTGATGATTCGCAAGAACT	6783
OY	1660	TTGATGCGATTAATTGCTGACAACTTCCACACAGCGACCCGATTTGGAAACGACATCG	1719
Db	6784	TTGATGCAATCATAGCAGACAACTTCMAAGCAAGGATCCCGTACTGGAGCGGATATTCG	6843
OY	1720	CGTGTGTGATTAAGCGCAGAGACGACGCTATGCGCATTTCCGCGTGTGATGATCCCTTAG	1779
Db	6844	CCTCGTTGACAAAGCCCAAGCGACGCTTATGGGGTTAACCGGGCTGATGATCTTTGGAAG	6903
OY	1780	ACTTAAAGTGTGACCAACCCGCTTTAGATTGTATGAGGCGGCGTTCCGCAATATTCAT	1839
Db	6904	ACCTGGGTTGGAACCACTACTCTCGACTTGATCGACTGCGCTTTGGAGAAATATCAT	6963
OY	1840	CTGTGCACTTACTCTACAGAGACGAGGTTTAAATTTGGCCATGATGAAATCCGATATGT	1899
Db	6964	CCACCCATCTGCCACGSGTACCCGTTTCAAAATTCGGGGCATATGAAATCCGAAATGT	7023
OY	1900	TCTTAAGCGTGTGTTGTCAACACTAGTCAATATATCATGATTTCTAGCAGAGTACTACGTG	1959
Db	7024	TCTTCACGCTCTTTGTCAACACAGTTCTGAATGTCTGTTATGCGCAGCAGAGTATTTGAGG	7083
OY	1960	AAAGGTTAACACGTCACCGTCGCGGCTCTATCCGGGAGATTAACHTAGTGAATGTGTG	2019
Db	7084	AGCGGCTTAAACGTCMAATGTGAGGATTTATCCGGGAGACAACTTATPACAGAG	7143
OY	2020	TGCTCTCCGACACTTGAATGCGGAGAGATGCGCATTGTGCTGACAACTGAGAAATTAATA	2079
Db	7144	TAGATATCGACAAAGAAATGTGCTGAGAGTGTGCACTGGCTCAACATGAGGTTAAGA	7203
OY	2080	TTATTTGATGCAATTATTTGTTACAAAGCACTTACTTTGTGGGGGATTTATCTGCTGG	2139
Db	7204	TCATTTGACGACGTATCGGCGAGAGACCACTTACTTCTGGGTGGATTCATCTTGGCAAG	7263

OY	2140	ACGAGTAAACGAGCAGACGCCGCGAGAGTCGACAGCCCTCTAAAAAGGCTTTTAAAGCTTG	2199
Db	7264	ATTGCGTTACTCCACAGCGTCGTGCGTGGGAGCCCTTGAAAAAGGCTGTTTAACTGG	7323
OY	2200	GAAGAACATTGECGAGTGCATGATATCCCAACACTGCGACCGCGCGGACATGCATATG	2259
Db	7324	GTAACACGGCTCCAGCCGAGCGATGAGCAAGACGAAGACAAAGACGGGCTCTGTGTAATG	7383
OY	2260	AAGCAATGCGATGGAAACGAATTGGAAATTACGACGAGTTAGTGAAGCCGTGAAATCCA	2319
Db	7384	AAACAAAGGCGCTGTTTAAAGTAGTATACAGACACTTAGCAGTGGCGGTGGCAATC	7443
OY	2320	GATACGAGTATATCTGCGACGGCCGTATCATCATCTCTGTCCACGTTAGCCGAAAGCG	2379
Db	7444	GGTATGAGAGTAGACACATCAACACTGTCTCTGTGGATTTGAACTTTTGGCCAGAGCA	7503
OY	2380	TTAAGAACTTCAAGAGCATAAAGAGGAGCCCAATCATCCTCTACGCGTGCATTAATAGG	2439
Db	7504	AAAGAGCATTTCAAGCCATCAGAGGGGAATTAAGCATCTCTACGGTGTCTTAAATAGT	7563
OY	2440	TGACGTAGTAGA-----CAGCGACCTTACCCACCGGACAGATGTTTCCA	2482
Db	7564	CAGCATAGTACATTTCACTGATCTAATATACCAACACACACACATCAATATAGAGATTTCT	7623
OY	2483	TACCTCAGCTGAACCTTCCACCAAGTTTACCCTTCAAAATCCGATGCTTACCGAGATCCA	2542
Db	7624	TTAACTATGCTCGCGCGCGCCCTTCCACGCCCCCACTGCTCATATGTGAGGCGCGGAGAA	7683
OY	2543	AAACCTCTTAAAGGCGCGCTGGAAGGCGGTTTCGGCCCCCGGTGGCTGCTCAATTCGAAGAT	2602
Db	7684	GGAGGACAGGCGGCCCCGATGCTGCGCCGCAATGGGCTTGTCCCAATCCAGCACTTGA	7743
OY	2603	CTTAGAGAGTTCGATAGTCAACTTGCATCTTCAAAACAC--GATCACTTAATCCGCGCGCAG	2660
Db	7744	CCACAGCGCTGATGCGCCTTAGTATTTGACAGAGGAATTAACCTCAAAACCCACAGCCAC	7803
OY	2661	GTCACACGGCCAAAGAAAGAAAGATGCTCTTAAAGCCAAACCTTACGCTTAAATAAGA	2720
Db	7804	GCCCCGCGCGCCGCAAAAGACGCGCCAAAGCAACCAACCGAAGCCGAAAGAACCA	7863
OY	2721	AGAGCGACGAAGCCAAAGGACGAAAC--GCAAGCTTAACAGAGGAAGCGAACACTGA	2777
Db	7864	AAACACAGGAGAAAGAAAGAAAGAACCTTGCAAAACCCAAACCCGAAAGAGACAGGTGA	7923
OY	2778	TGTTGATGAAGTTGAGATCGGACAAAGACATTTCCGATC--ATGCTGAAGCGGCAAGTGA	2834
Db	7924	TGGCACTTAAGTTGAGAGCGCGACGACGATGTTTCACGTCATAAATATGAGACGAGATGTCA	7983
OY	2835	ATGGAATATGCTCGCTGTGTGAGAGAAAGCTGATGAACCACTTCCAGCTTGAAAGAAAA	2894
Db	7984	TCGGGACACGCACTGGCATGGAAGAAAGGTAATGAACCACTTCAAGTGAAGAAACTGA	8043
OY	2895	TTGATTAATGAGCAATTAGCGGCGGTGAATTTGAAGAAAGCTAGCATGTACGACTTGGAGT	2954
Db	8044	TTGACCAACCTGTGCTATCAAAAGCTCAAAATTTCAACAAAGTGTGACGACTAAGACATGAGT	8103
OY	2955	ACGCGACAGTTCCTCCCAAAATGAAATCAACACAGCTGACAGTACACCGACGCAAAACAC	3014
Db	8104	TCGCAACAGTTGCCGGTCAACATGAAAGTAGGCGTTCACTTACCAACGATGAAGACCTTG	8163
OY	3015	CGGGCTTCTACACTGGCACCAAGCGCGCACTGATGAGAAATGGGAGATTTTACCGTAC	3074
Db	8164	AAAGGTTCTAACATGTGSCACACAGAGCGGTGACGATGTAGTGAGGACAGATTTTACCATCC	8223
OY	3075	CGAAGAGATGCGCGGGAAGGCGACAGCGGAAAGCCTGATCTCTGACACACAGAGGCAAG	3134
Db	8224	CCCGCGAGTAGAGAGGCGAGAGGAGACAGTGTGTGTCGATTAATGATATACCTACAGCGCGG	8283
OY	3135	TTGTGGCTATTGTTCTTAGAGAGTGCAAATAGAGGCAAGCTTACGGGCTTTTACGTGATCA	3194
Db	8284	TTTGTCGGAATGTCCTCGAGGGGCTGATAGAGGAACAGAACCGCCCTTTTCGGTGTCTCA	8343
OY	3195	CTTGGAAACGAAAGGGGTGACCAATTAGGAGTACCCCGAAGGTTTGAACCGGTGT--	3251

8244 CCTGGAATAGCAAAAGGGAAGCAATCAAGACACCCCGAAGGGAAGAAAGTGTCTG 8403
3352 -----CACTAGTTACAGCGCTATGCGTCTTTGGAATGTCAAGTTCCCATGCGACAAAC 3305
8404 CTGCACACCTGGTCAACCGGCCATGTGCTTGTGGAAGAGTGAAGTCTCCATGCAATCGCC 8463
3306 CACCGGTGTGTATTCACTGACGCGAAGAACGAACTGCAAGTGTGGAAGAAAGTGTG 3365
8464 CGCCCAATGCTTACACCGCGCAACATCAAGAGCTCTTGACATCTCTGAAAGAAAGTGA 8523
3366 ACAATCCAAATTACAGACAGTGTGGAAGAGTCTTGAAAT---GTCCATCAGCGCGGC 3422
8524 ACCACGAGGCTTACAGACCTGTCTCAACGCCATATTCGGTGGGAGTGTGCGGCAAGAA 8583
3423 CCAAAAGCAATTAACGATACCTTACATGACCAAGTCTTACTGGGGTCTTGCCCT 3482
8584 GTAAAAAGAGGTCACTGACACACTTAACTTGAACCGCCGTACTTGGGCAATGTCTGT 8643
3483 ATTGACAGACTCAACGCGGTGTTCAGCGCCCAATAAAAATTGAAGAGTGTGGGACGAT 3542
8644 ACTGTCACTACTGAAACCGTGTCTTTAGCCGATTAGATGAGACAGGTCTGGAGATGAAG 8703
3543 CTGATGATGATCGAATTAGAAATCGAGGTCTGCGCACAATTGGCTACATCAGGACGCA 3602
8704 CGAGCAACAACCATACGATACAGACTTCGCCCAAGTTTGAATAGACCAAAAGCGGAG 8763
3603 CTGGGATGTACCAAAATTCGTTTACATGTCTTTGACACAGACCATGACATCAAGGAAG 3662
8764 CAGAGAGCTCAAAATAGATACCGCTACATGTGCTGACAGAGATCATCTGTCAAAAGAG 8823
3663 ACAGATAGAGAAATAGCTATCAGACATCTGAGACCTGCGGTGTGTTGGCCCAAG 3722
8824 GCACCAATGATGACATCAAGATCAGACCTCAGACCGGTGTGAAGGCTTAGCTACAAAG 8883
3723 GGTACTTCTGTAGCTCAATGTCTCTCAGGTGACAGTGAACCGTACGATCAAGAGCG 3782
8884 GATCTTTCTCTCGCCAAAGTGTCTCCAGGGGAAGGATTAAGCATAGAGGATA 8943
3783 GAGCATCTGAGAAATTCATGACCGGTGAGAAAAAGATAGAGAAAGTTGTGGTAGAG 3842
8944 GCAACTCAGCAACGTATGACAAATGCGCCGCAAGATTAACCAAAATTCGTGGACCGG 9003
3843 AGAGTACTTGTGCCACCGGTCAATGGAAGGTGTAAGTGCACGCTTAACGATCACT 3902
9004 AAAAATATGACTTACTCTCCCTTACCGTAAAGAGATTCCTTGACAGTGTACACCGTTC 9063
3903 TGAAGAGACGTCTGCGCGGTACATTAACCATGACAGGCGCAGGCCACACCGCTATAGT 3962
9064 TGAAGAAACAAACCGCGGTACATCACTATGACAGGCGCGGACCGCATATACAT 9123
3963 CCTATCTGAGGAAGCGTCAGCGCAAGTGTACATTAAACACTTCTGGCAAGAACGTCA 4022
9124 CCTATCTGAGGAATCATCAGGAAAGTTTACGGAAACCACTCCCGGAAAGAACATTA 9183
4023 CCTACGATGTAAAGTGTGGGACATACAGACAGGATGTGAGACCGGGAAGCAAGATGA 4082
9184 GGTACAGTGTAGAGTGTGGGCAATTAAGACCGGAACCGTTTACGACCGGTACCGAAATCA 9243
4083 ACAGCTGCACTTAAAGCAAAACAGTGCATTTGCTTACAGAGCGAACCAAGAAAGGTCT 4142
9244 CGGCTGACCGGCATCAAGCAGTGTGTGCTTAAAGAGCGAACAAAGAGTGGCT 9303
4143 TCAACTGCGCGGATTTATTGGCACACAGACCACTCAGTGCAGAGTAAATTGCAATTC 4202
9304 TCAACTGCGCGGACTCGATCAGACACGCGAACCAACGCGCCCAAGGAAATTTGATTTGC 9363
4203 CATTCGCTTGACACCGCAGATGTGCGCGGTTCGTTAGCTCAACGCTACAGTACGA 4262
9364 CTTTCAAGCTATCCGAGTACCTGATGTGTCTGTGCGCAGCGCGCAACGATGATAC 9423
4263 AGTGTTCAAAGGATCAACCTCCACCTGACTGCAATGCAACCAATTTGCTGACAAAG 4322

9424 ACGGCTTAAACACATCAGGCTCCCAATTAGACACAGACCAATCTGACATTTGCTCACCA 9483
4323 GAAATTTGGGGCTGTGAGAGAGCAACAGACAGAAATGAAATTAAGGGTTATACCAAGA 4382
9484 GGAAGTAAAGGCAAAACCGGAACCAACACTGATATATATGGAACACGGTTAGAA 9543
4383 ATTTTCTGTGGGCGAAGAGGCTGAGTACGATGTGGGTAAACATGAACCACTCAGAG 4442
9544 ACTTCAACGTGTGACCGAGATGTGGCTGAAATACATATGGGGAAATCAAGAACATTAAGGG 9603
4443 TGTGGCCCAAGAGTGTGACACAGGAGCCCAACATGATGTGGCGCATGATCATTC 4502
9604 TCTATGCCAAAGTGTGACACAGAGACCTTCAGATGTGCCCAAGAAATATGATAGC 9663
4503 ACTATTAATTCAGGCACTGATCAGCTGATCTGATGTTGTTGTTGTTGCTTGTGCTA 4562
9664 ATTACTATCATCGCATCTGTGTGTACACCATTTAGCGGTGATCAGCTGTGTGGCA 9723
4563 TCTGTAGGCACTGATCATCAGACAGCTTGACATGCGCAAAAGCAAGAGACTGCTGA 4622
9724 TGATGATTTGGGTAACTGTGGACATTAATGTGCTGTAAAGGCGCGGTGAGTGGCTGA 9783
4623 CGCCATACGGCTTTGACACCGAAGCAAGGTATCCACAGATTAAGCGGTTTGTGCTGA 4682
9784 CGCCATATGCTGTGGCCCAAGTCCGTGATTCACATTTGCTGCACTTTTGTGCTGTG 9843
4683 TTGGCCCAACCAACGCTGAACATTTGAGAACTTTGAACCATCTGTGTTTGAACACC 4742
9844 TTAGTGTGGCTAAAGTGTGAACATTCACCGAGACATGATGTTATTTGTGCAAGCC 9903
4743 AACGTTTCTGTGGGCAAGTGTGATCTCTGTGGACGCGTTGTTATCTGTTCCGCT 4802
9904 AGCGTTCTTGTGGGTCAAGTGTATACCTCTGTGCGGTGTGTCTTAATGTGCGT 9963
4803 GCTTTATGCTGATGCTTTTATTTATGTTGTGACAGCGTGTGCTGTGGAAAGTATAGC 4862
9964 GTTGTCTATGCTGCTCTCTTTTATGTTGTGCTGCGGCTACCTGTGCAAGGTATAGC 10023
4863 CCTTGAACATGCGACCACTGTGCAATGTCCGGGATTCGGTATTAAGGCGTTGTGCG 4922
10024 CTTAGGAACATGCAACCTGTTCAAAATGTGCCACAGATACGATTAAGCACTTGTG 10083
4923 AACCGAGGTTTACGCGCACTTAACCTGAGATCAGGTGTCTACATCGGAATTAACAC 4982
10084 AAAGGAGGGGTACGCCCGCTCAATTTGGAGATTATGTCATGTCTCCGAGGTTTGC 10143
4983 CTTCACTTAACAGAGTGTGACCTGTGCAATTCACACAGTATTTCTTACACCAAG 5042
10144 CTTCCACCAACCAAGAGTATTAATCTGCAAAATTCACACCTGTGTGCTCCCTCAAG 10203
5043 TTAATGTGTGGGTCTCGAGTGAAGGATCTCAAGAGGCGGATTAACATGCGCGG 5102
10204 TCAGATCTGTGGCTCTTGAATGTCAAGCCGCGCTACAGCACTATACCTGCAAG 10263
5103 TTTTGGCGGTGTGTAACCTTTTCAATGTGGGAGGCGCAATGCTTCTGTGACAGTGA 5162
10264 TCTTTGAGGGGTGTACCCCTTATGTGGGAGAGACCAATGTTTGGACAGTGA 10323
5163 ACACACAACTGATGAGGCGTGTGAGTTCGCTCCAGACTGACATATATGATACCGCAG 5222
10324 ACAGCCAGATGAGGAGGCGTACGTAATTTGTCAGTAAATTTGCGCACTGACCAACCGC 10383
5223 TCGACCTAAAGTTCAACAGCTGTCTGAAAGTGTGCGGTATGATTAAGGCAACA 5282
10384 AGGGATTAAGGTCAATCTGCGGAGTGAAGAGTGTGCGTATGTATCGGGAACA 10443
5283 CCACCGGCACTGTGATACGTTTGTCAATGGCGTCAAGCGATTTCTCAAGGACCTGA 5342
10444 CTAACAGTTTCTTATGATGTGTGATGTAACGAGTCAACACAGAAACGCTTAABAACCTGA 10503
5343 AGGTATAGCAGGGCGCATATCAGCGCTTTTTCACCTTTGACATTAAGGTGTCTATCA 5402
10504 AAGTCATAGCTGACCAATTTCAAGCATTTGTTTACACCATTCGATCAAGAGTGTATCA 10563

```

QY 5403 GAAAGGGGCTTGTACACTACGACTTCCCTGAGTATGAGCTATGAACCGAGAGCGT 5462
DB 10564 ATCCGGGGCTGGTGTACAACTATGACTTCCGGAATACGAGCGATGAACCGAGAGCGT 10623
QY 5463 TCGCGGATTTTAAAGCATCTTGGCTTGTATGATCAGACATAGTACCGGCTGACATAC 5522
DB 10624 TTGGAGACATTCAGCTACCTCTTGTAGTACGAAAGACTCATCCGACACAGACATTA 10683
QY 5523 GGCTGCTGAGGCTTCTGTCAAGAACATCAGCTCCCTTACACCCAGAGATATCAGGT 5582
DB 10684 GGCTACTCAAGCTTCCGCAAGAACGTGATGCTCCGACACGAGCGGATCTGAGT 10743
QY 5583 ATGAATGTGAAAGAACATCAGACGACGCCCTGCAAAAACAGACCATTTGATGTA 5642
DB 10744 TCGAATGTGAAAGAACATCAGACGCCCTGCAAAAACAGACCATTTGATGTA 10803
QY 5643 AAATTGAAGTGAAGCTCTGCGAGCGTCTAAGCTGTGCTTACGGGACATCCTATCTGA 5702
DB 10804 AGATTGAGTCAATCCGCTTTCAGAGCGGTGAGCTGCTCATACGGGAACATTTCCATTTCTA 10863
QY 5703 TTGACATCCCTGATGAGCTTTTGTGAGATCATCAGATCACCACAAATTTTGAAGTTA 5762
DB 10864 TTGACATCCCGAAGCGCTGCTTATCAGAGACATCAGATGACACACTGCTCAACAGTCA 10923
QY 5763 GCTGCAAGTACAGACGCTCATTTATTTGAGAGACTTGTGTGTTCTTAACATTAAGT 5822
DB 10924 AATGTGATGTCACTAGTATGATCACTTATTCAGCGGACTTGGAGAGATGCTACCTGCACT 10983
QY 5823 ACAAACTGACAGGAGGAGGAGCATTTGTCAGATTCTCCACTCCAGACAGCTGTTTTGA 5882
DB 10984 ATGTATCCGACCGCGAAGGAGCATGCTGTCATTCGATTCGAGCAACAGAACCTTC 11043
QY 5883 AGGAAGCCACACATGTGATGCTGCGGTAGGACGATTAACATTAATTTTACATTCGA 5942
DB 11044 AAGAGTCGACAGTTCATGTCTGTGAGAAAGAGCGGTGACATTAACCTTACACACGCGCA 11103
QY 5943 GCCCAAGCAAAATTTATAGTTTGGCTATGCGGCAAGAAAGTCCACTCAATGCTGAT 6002
DB 11104 GCCCAAGCGGAATCTTATGATGCTGTGTGTGTAAGAGACATGCAATGCAAGAT 11163
QY 6003 GTAAACACCGGCGGACCATTAATTTGAGAAACACATTAAGTGCACCAAGATTTCCAG 6062
DB 11164 GCAAAACACAGCTGATCATATCGTAGACACCCCGCAAAAATGATCCAAAGATTCAG 11223
QY 6063 CGGCAAGTTCCAAACATTTGGAAGTGTGCTGCTGCTGCTGTTGGGGAGCATATCCC 6122
DB 11224 CCGCATCTCAAAAACCTTCATGAGAGTGTGCTGTTGCTTTCCGCGCGCTGCTGCGC 11283
QY 6123 TCATTGTTGATGAGCTTATAGTGTGTGCTGAGAGCTTATGCTTAAACACAGTATG 6182
DB 11284 TATTAAATTAATGAGCTTATGATTTTGTCTTGCGAGCATGATGCTGATACACAGAAAGT 11343
QY 6183 GACTAGGCGCGGACACTGA-CATAGCGGTAAACTGATGATCTTCCGAGAGCGTGT 6241
DB 11344 GACCGCTAGCGCCCATGACCCGACACAGAAAACCTGATGATCTTCCGAGAACTGAT 11403
QY 6242 GCATATGCCACGCGCGCT-----TGAAGT 6268
DB 11404 GCATATGATGATGAGCTGTGATATTAATAGATCCCGCTTACCGCGGCAATATGCAACAC 11463
QY 6269 AAAATCGATGATTTTCCGAGAGACAGCATTAATGCTGCGAGT 6318
DB 11464 AAAATCGATGATTTTCCGAGAGAGCGCATGATATGCTGCGAGT 11513

```

```

RESULT 9
ABX81526
ID ABX81526 standard; cDNA; 11740 BP.
AC ABX81526;
XX
XX 23-APR-2003 (first entry)

```

```

XX Sindbis virus cDNA #3.
DE Eukaryotic layered vector initiation system; cytopathic effect; gene; ss;
XX alphaviral nonstructural protein 2; RNA vector replicon; alphavirus RNA;
XX macromolecular synthesis; alphavirus RNA polymerase recognition sequence;
XX 3' polyadenylate tract; cellular macromolecular synthesis; cell death;
XX Sindbis virus.
OS Sindbis virus.
XX
XX US6465634-B1.
PN
XX
XX 15-OCT-2002.
PD
XX
XX 08-OCT-1999; 99US-0415900.
PF
XX
XX 06-OCT-1997; 97US-0944645.
PR
XX 05-APR-1996; 96US-0628594.
PR 24-JUN-1996; 96US-0668953.
PR 12-JUL-1996; 96US-0679640.
PR 04-APR-1997; 97US-0833148.
XX
XX (CHIR ) CHIRON CORP.
PA (UNIV ) UNIV WASHINGTON.
XX
XX Dubensky TM, Pojo JM, Belli BA, Schlesinger S, Dryga SA, Frolov I;
XX WPI; 2003-147073/14.
DR
XX
XX Eukaryotic layered vector initiation system, for gene therapy, has
XX alphaviral nonstructural protein gene having mutant nonstructural
XX protein 2 gene, which reduces host-cell directed macromolecular
XX synthesis.
XX
XX Example 1; Fig 8A-8E; 161pp; English.
XX
XX The invention relates to a eukaryotic layered vector initiation system,
XX comprising a nucleic acid sequence encoding all four alphaviral
XX nonstructural proteins and including an altered sequence encoding for
XX nonstructural protein 2, such that when the altered sequence is operably
XX incorporated into an RNA vector replicon, the time required to reach 50%
XX inhibition of cellular macromolecular synthesis in cells is increased, as
XX compared to an RNA vector replicon having a wild-type alphavirus
XX nonstructural protein 2. The initiation system comprises a 5' promoter
XX which directs synthesis of alphavirus RNA in vivo from cDNA, a 5'
XX sequence which directs transcription of alphavirus RNA, a nucleic acid
XX sequence which operably encodes all four alphaviral nonstructural
XX proteins, an alphavirus RNA polymerase recognition sequence and a 3'
XX polyadenylate tract. The eukaryotic layered vector initiation system is
XX useful for stimulating an immune response within a vertebrate, for
XX protein expression and gene therapy. The system exhibits reduced, delayed
XX or no inhibition of cellular macromolecular synthesis, thus permitting
XX its use for protein expression and gene therapy with reduced, delayed or
XX no development of cytopathic effects or cell death. This sequence
XX represents Sindbis virus cDNA, used in the scope of the invention.
XX
XX Sequence 11740 BP; 3345 A; 3049 C; 2908 G; 2438 T; 0 other;
SQ

```

```

Query Match 36.7%; Score 2380.4; DB 25; Length 11740;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3693; Conservative 0; Mismatches 1986; Indels 41; Gaps 8;

```

```

QY 580 ACTGACGTATGACGCGGAGCGGTATTTTCTATCGGAACAGCCAGGTACCTTC 639
DB 5746 ACTGACGTATGACGCGGAGCGGTATTTTCTATCGGAACAGCCAGGTACCTTC 5805
QY 640 AACAGAAATCAGTATGATTAACAGAACTATATTTGATGCGCGCGCATG 699
DB 5806 AAAAGAGTCCGTTTGGAGAACGCTTACGAACGACCTTGAAGCGCAATGCTCG 5865
QY 700 AGAAGTATACGCCCGCGCTGATCTCGAAAGAGAAAATTTTACAGAAAGAACTGC 759

```

Dh 5866 AAGAAATTCATGCCCCGGTGTGCAACGTCGAAAGAGAACTCAAACTCAGTACC 5925
Qy 760 AATTATGCGCTCTGGAAGAAATAGAACAGTATCATCAGAAAAATGAAAAATATGA 819
Db 5926 AGATGATGCCCAACGGAAGCCAAAGATAGTACCACTCTGAAAGTAAATAATCGA 5985
Qy 820 AAGCAATTAACAGCGAGGAGCTCATTTCTGATTTGGGACATATCTATCATCAGAAATGA 879
Db 5986 AAGCATTAACCACTGAGCGACTACTGTGAGAGACTACAGCTATTAACCTCTCCACAG--- 6042
Qy 880 ATCTGTGAGTGTAAACAGTCAATATCTGTACCAATCTACTCTGCAACGGTAAATTA 939
Db 6043 ATCAGCAGAAATGCTATTAAGATCACTATCCGAAACATGTACTCTCCAGTACGGTACCGG 6102
Qy 940 ACAGSTTACATCTGAGAGAGTCCCGGTAAAGGTGAACTTATAGTTATCCAAAGAAAT 999
Db 6103 CGAACTCTCGAATCCACAGTTCGCTGTAGTCTGTACAACTATCTGTGATAGAACT 6162
Qy 1000 ACCCTACAGTACGAGTATATGTATTAACAGATGAATACGATGCGTATCTGACATGGTGG 1059
Db 6163 ATCCGACAGTACGATCTTTATCAGATTAAGTACGAGTACGATGCTTACTTGATATGGTAG 6222
Qy 1060 ACGGCGATCGTGTCTGTAGATACAGCACTTTTGTCCGGCTAAACCTGAAGACTACC 1119
Db 6223 ACGGACAGTGTGCTGCTGTGATCTGCAACCTTCTGCCCCGCTAAGCTTAGAAGTTACC 6282
Qy 1120 CAAGAAGCATAGCTATTTGAGAGCAAGATTAAGATCAGCCGTCCCATGCGCTATACGA 1179
Db 6283 CGAAAAAATGAGATATAGAGCCCGCAATATCCGAGTGTCCATACAGCATGACGA 6342
Qy 1180 ATACATTACAAATATGATTTGGCTGACGCTACTAAAGAAATTGCAACCTTACCCAAATGC 1239
Db 6343 AACAGCTACAAATATGTGCTATTTGCCGCACTAAAGAAATTTGCAACCTCAGCAGATGC 6402
Qy 1240 GAGATTAATCTGTCTTATGATTCGCGCGCATTAATGTTGATTTTCAAGAAATACGAT 1299
Db 6403 GTGAATCGCCAACTGAGTACGAGTACGACATTCATATGTGAAATGCTTTCGAAATATGAT 6462
Qy 1300 GCATGATGATGATCTGGGATACCTTTGGCGAATACCTATTCGGCTAATCAAGAAACG 1359
Db 6463 GTATGACGAGTATTTGGAGAGTTCGCTCGAAGCCAAATTAAGATTAACCTAAGTTTG 6522
Qy 1360 TTACGCAATATGTGACAAAGCTGAAAGGCGGAAAGCAGACGATTTGTTSCGAATCTC 1419
Db 6523 TCACCGCATATGTAGTACGTAAGAGCCCTTAAGGCGCGCAGCTATTTTCAAAAGAGCT 6582
Qy 1420 ATTAATCTAAACCGTTCGAGAGATACCAATGATCAATTCGTATGATGATTAAGAGAG 1479
Db 6583 ATTAATTTGTCCTCATAGAGAGTCTGATGATGATTCGTATGATGATGAAAGAG 6642
Qy 1480 ATGTCAAGATTAATCTCCGGGACGAAACCTACAGAGAGCGGCTTAAGTGCAGGTTATTC 1539
Db 6643 ACGTGAAGATTAACACAGGACGAAACACACAGAAAGAAACGAAAGTAAAGTGAATAC 6702
Qy 1540 AAGGTGAGATCCCTTGTACCGCTTACCTTTGCGGAGTCCATTCGGGATTTAGTCGCTA 1599
Db 6703 AAGCCGAGAAACCCCTGCGAGCTGCTTACTTATGCGGATTCACCGGAAATTAATGCTGA 6762
Qy 1600 GACTGAATGCGGTCTTGTGCAAAATATCACTACTCTTTCGACATGTCAAGGGAAGTT 1659
Db 6763 GGGCTTAAGCGCGCTGTGCTTCAAAACATTCACAGCTTTTGAATGTCGCGGAGGATT 6822
Qy 1660 TTGATGCGATTAATGTGTAATTTCCACACGCGGACCCAGTATTTGAAACGAGCATCG 1719
Db 6823 TTGATGCAATCATGACAGAACTTCAAGCAAGCGACCCGGTACTGAGACGGAATTCG 6882
Qy 1720 CGTGTGTTGTAAGAGGAAGACGATCGCATTTGCGCGTGAATGATTCCTTGAAG 1779
Db 6883 CATATTCGACAAAGCCAGAACACGATATGCGCTTAACCGGTGATGATCTTTGAGAG 6942
Qy 1780 ACTTAGTGTGACCAACCGCTCTTAATTTGATAGAGCGCGTTCGGCAATATCAT 1839
Db 6943 ACCTGGGTGTGATCAACCACTACTGCTTATGATGAGTGGCTTTGGAGAAATATCAT 7002

Qy 1840 CTGAGCACTTACCTACGAGAAACGAGTTTAAATTTGGTGCATGATGAAATCCGGTATGT 1899
Db 7003 CCACCACTTACTACCGGATCTCTGTTTAAATTTGGGCGGATGATGAAATCCGGAATGT 7062
Qy 1900 TCTTAACGCTGTTTGTCAACACACTACTGATCAATATGATGATTTGAGAGATTAACGTG 1959
Db 7063 TCTTCACTATTTTGTCAACACAGTTTGAATGTGTTATTCGACAGAGTACTTAAG 7122
Qy 1960 AACGTTAAACAGTCAAGCGTGGCGGCTCTATCGCGGACGATTAACATAGTGAATGTG 2019
Db 7123 AGCGCTTAAACGCTCAGATGTGACAGCTTCAATTTGCGACAGCAACATATCATATGAG 7182
Qy 2020 TCGTCTCGACACCTTGATGGGAGAGATGCGGCACCTTGAGCTGAACATGAAATTAATA 2079
Db 7183 TAGTATCTGACAAAGAAATGCTGAGAGGTGCGCCACTGCTCAATGAGAGTTAAGA 7242
Qy 2080 TTATTTAGTGAATTAATTTGATATCAAGACACCTTATCTGTGGGGAATTAATCTGTGTG 2139
Db 7243 TCATGACGAGCATCGTGAAGAGACCACTTACTCTGCGGGGATTAATCTTTCGAAG 7302
Qy 2140 ACCAGATTAACAGCACGCTGCAAGATCGCAACCTCTTAAAGGCTTTTAAAGCTTG 2199
Db 7303 ATTCGTTACTTCCACAGCGTGTGCGCGGTGCGGATCCCTGAAAGGCTGTTAAGTTGG 7362
Qy 2200 GAAACCAATGCGAGTGAATATCCCAAGCTGCGACCGCGCGGCACTGCATGATG 2259
Db 7363 GTTAACCGCTCCCGCGGACGACGAGCAAGACAGAAAGACGCGCTCTGCTAGATG 7422
Qy 2260 AAGCAATGCGATGGAACAGAAATGGAATTAACGACGATGATGAGAGCGGTAGAAATCA 2319
Db 7423 AAACAAAGGCGTGTATTAAGTATGATGATTAACGACCTTATAGCAGTGGCGTGAACACC 7482
Qy 2320 GATACGATATCATCTGAGCGCTGATCATCACTGTCTGTCCAGCTTACCGAAAGCG 2379
Db 7483 GGTATGAGGTAGCAATATTAACACTGTCTCTATGAGATTAAGAACTTTTGGCCAGAGCA 7542
Qy 2380 TTAAAGACTTCAAGAGCATTAAGAGGAGGCCAATCACCTCTACGCTGACCTTAATAGG 2439
Db 7543 AAGAGCATTTCCAGCCATCAAGGAGGAAATTAAGCATCTCTACGTTGTGCTTAATATGT 7602
Qy 2440 TGAAGTATGACACGACGCTAC-----CCACCGGCAAGTGTTCATATCC 2486
Db 7603 CAGCATATGATCATTTATCTGATCAATATATCAACACACACCATGATTAAGAGATTC 7662
Qy 2487 CTCAGTGAATTTTCCACAGCTTATACCTTCAAAATCCGATGCTTACCGAGATCAAAAC 2546
Db 7663 TTAAATGCTCTGCGCGCGCGCCCTTCCCGGCCCACTGCGATGTGAGGCGCGGAGAA 7722
Qy 2547 CTCTAGGCGCGCTGAGGCGGTTTGGGCGCGCGCTGCTCAATGGAATCTTTA 2606
Db 7723 GAGAGCAGGCGGCGCCGATGCTGCGCAAGGAGTGGCTTCTCAATTCAGACATGTA 7782
Qy 2607 GAGAGTGAATGATCACTTGAATTTTCAACAAAGATCACTTAATCCGCGCAGTCA- 2665
Db 7783 CCAAGCGCTCAAGGCGCTGATGATGACAGGAACTTAAGACCTCAACCCCAAGTCCAC 7842
Qy 2666 -----CGCCAAAGAAAGAAAGTCTCTTAAGCCAAACCTTACAGCTTAAAGAA 2720
Db 7843 GCCCGCACCGCGCAGAAAGAGCGCGCCCAAGCAACACCGAAAGGAAACAA 7902
Qy 2721 AGAAGCAGCAAGCCAAAGAGACGAAAC---GCAAGCTTAAACAGGAAACGAAACGTA 2777
Db 7903 AAACGCAAGAAAGAAAGAAAGACGAAACCTGCAAAACCGGAAAGAGACAGCGCA 7962
Qy 2778 TGTGTATGAAGTTGAGTGCAGACAAAGATTTTCGATCATG---CTGAACGCGCAAGTGA 2834
Db 7963 TGGCACTTAATGTTGAGGCGGACAGATTTTTCAGCTCAAGAAACGAGAGAGATGCA 8022
Qy 2835 ATGATATGCTGCTGTGAGAGAAAGCTGATGAACCACTTCACAGTTGAAGAAATA 2894
Db 8023 TCGGACGACGACTGGCCATGGAAGAAAGTAAATGAAACCTTTCACAGTGAAGAAACCA 8082

QY 2895 TTGATATAGCAATTCAGCGCCGTGAAATTGAGAGGCTAGCATGTACGACTTGAGT 2954
Db 8083 TCGACCACTGTGTGATCAAGCTCAATTTTACCAGTGGTACGATACGATGGAGT 8142
QY 2955 ACGGGAAGTTCCTCCAGAACATGAAATCAGACCGCTGCACTGACCGGCAAAACAC 3014
Db 8143 TCGACATGTTCGATCAACATGAGAGGCAATTACCTACACAGTGAACCGCCG 8202
QY 3015 CGGGCTTTACAACTGGGACCAAGCGGAGTCCAGATGAGAAATGGAGATTACCGTAC 3074
Db 8203 AAGGATTTATTAATCTGGACACCGAGCGGTGCAATATGAGAGTATGATTTACATCC 8262
QY 3075 CGAAGAGATGGGCGGAGAAAGCGACAGCGGAGACCGATCTGTGACAAACAGAGCAGAG 3134
Db 8263 CTCGCGAGTAGAGAGGAGAGAGACAGGGGTGGTCCGATCATGATTAATCCGGTCCGG 8322
QY 3135 TTGTGGCTATTTGTTTAGAGAGGTGCAATGAGGGGCAAGGCTTCGAGTGTCA 3194
Db 8323 TTGTGCGATAGTCTCGGTGGCGCTGATGAGAGAACAGAACTCCCTTCGGTGTCA 8382
QY 3195 CTGGGAACGAGAAAGGGGACCATTTAGGAGATACCCCGAAGGTTCTGAAACCGTGT--- 3251
Db 8383 CTGTGAATAGTAAAGGAGAACATTTAAGACGACCCGAGAGGACAGAGAGTGTCCG 8442
QY 3252 -----CACTAGTTACAGCGCTATGCGTCTTTCGATGTCACTTCCCATGCGAAC 3305
Db 8443 CAGACCACTGTGTACCGGCAATGTGTTGCTCGGAAATGTGAGCTTTCCATGCGACCGC 8502
QY 3306 CACCGGTGTCTATTCTACGTAGCGGCAAGAACACTGACGTGTCTCGAAGAAAGCTGC 3365
Db 8503 CGCCCACTGTCTATACCGCGCAACCTTCAGAGCCCTGACATCTTTGAGAGAAAGCTGA 8562
QY 3366 ACAATCAAAATTAAGACAGCGCTGTGAGAAAGTGTGAAATGTGCATCAC---GCCGCG 3422
Db 8563 ACCTAGAGCGCTACATACCTGCTCAATTCCTATTTGGCGGATGCTGTGCAAGAA 8622
QY 3423 CCAAAAGAGATTAACCGATGACTTACACTGACCACTGACCTGAGGCTTCTGCGCT 3482
Db 8623 GCAAAAGAGGCTATTAAGACCTTTACCTGACAGCCCTCACTTTGGGCAATGCTGT 8682
QY 3483 ATTGCAGACTCAAGCGCGTGTTCAGGCCCAATTAATAATTGAGAACTGTGGGCAAT 3542
Db 8683 ACTGCACATATCTGTACCGTGTTCAGCCCTGTATGATCGAGAGGCTGTGGGAGCAAG 8742
QY 3543 CTGATGATGATTCGATTAAGAAATCCAGGTCTGGGCAATTCGGCTACATCAAGAGCA 3602
Db 8743 CGGAGATTAACCACTACGATACAGACTTCGCGCAAGTTGATGACGACCAAGCGAG 8802
QY 3603 CTGCGATGTACCAAAATTCGTTACATGCTTTTGACCAAGACCATGACATCAAGAG 3662
Db 8803 CAGCAAGGCAAAAGATACCGCTACATGTGCTTAAGCAGAGATCAACCGTTAAAGAG 8862
QY 3663 ACAGTATGAGAAATAGCTATCAGACATCTGGAACCTCGCGTGTCTTTGGCCAAAG 3722
Db 8863 GCACCATGATGACATCAAGATTAGCACCTCAGAGCCGTGTGAGAGGCTATGCTCAAG 8922
QY 3723 GGATCTTCTGTAGCTCAATGTCTCTCAGGTGAGAGGTAACTCTGATACAGAGC 3782
Db 8923 GATATTTCTCTCGCAAAATGCTCTCAAGGGGAGAGGTAAAGGTATGACATAGTA 8982
QY 3783 GAGCATCTGAGAAATTCATGACCGGTGAGAAATGATCAGAGAGAGTTGTCTGTAG 3842
Db 8983 GCACTGCAACGTCATGTACATGCGCCGGAATTAACCAAAATTCGTGGGACGG 9042
QY 3843 AGGAGTACTTGTTCACACCGCTGCATGAGAAAGCTGTAAAGTCCAGCTTTACATCACT 3902
Db 9043 AAAATATATGATCTACCTCCGTTACGGTAAAAAAATTTCTTGACAGGTACGACCTGC 9102
QY 3903 TGAAGGAGACGCTGCGCGGTATCATACATGCAAGGCGGACCAACCGGTATTAAT 3962
Db 9103 TGAAGGAGAACATGCGAGCTACATCACTATGCAAGGCGGACCGGCTTTATTAAT 9162
QY 3963 CCTATCTGAGGAGACGTCAGGCGAGGTATCAATTAAACCACTTCTGCAAGAGCTCA 4022

Db 9163 CCTACTGGAAGATCATCAGGAGAAAGTTTACGCAAAAGCCGCTATCGGAGAAACATTA 9222
QY 4023 CCTAGAAATGTAATGTGTGCGCATTAACAGACAGATATGTGAGACCGGAAAGAGATGA 4082
Db 9223 GGTATGAGTGAAGGTGCGGCACTAACAGACCGGAAACGTTTTCGACCGGCAACCAATCA 9282
QY 4083 ACGGCTGACTTAAAGCAAAACAGTGCATTCGCTCAAGAGGAGCAACCAAGAAATGGGTCT 4142
Db 9283 CTGTTGACACCGCATCAAGAGATGCTGCTGCTTATAGAGGACCAAAAGAGTGGTCT 9342
QY 4143 TCAACTGCGCGGATCTTATTAAGGACACAGACCACTCAGTGCAGAGTAATTTGACATTC 4202
Db 9343 TCAACTCACCGGACTTGTATCAGACATGACGACACACCGCCCAAGGAAATTTGACTTGC 9402
QY 4203 CATTCGCTTACACCGACAGTGTCCCGGTTCGCTTACGCTACACCGCTACAGTACAGA 4262
Db 9403 CTTTCAAGTTATCCCGGTATCTGCAATGTCTCGTTGCGCAGCGCGCAATGTAATAC 9462
QY 4263 AGTGTTCAAAGGCTACACCTCCACCTGACTGCATGCAATGCGAACCAATTGCTGACAGCA 4322
Db 9463 ATGCTTTAAACATCACTACGCTTCAATTAATGATACAGACCACTTGCATTTGCTACACCA 9522
QY 4323 GAAATTTGGGCTGGAGCAGACCGACACGAGAAATGATTAACAGGTCTACATCAGAGA 4382
Db 9523 GAGACTAGGGGCAAAACCGGAAACCAACCACTGAATGATGCTGGAAGAGAGCTGACAGA 9582
QY 4383 ATTTTCTGTGGGCGAGAAAGGCTGTGAGTACGTTAGGGGTAAACATGACATCAAGTCAAG 4442
Db 9583 ACTTCAACGCTGACCGAATGCTGTGAAATCATATGAGGAAATCATAGCCAGTGAAGG 9642
QY 4443 TCTGGAGCCAGAGTGGGACACAGGCGACCAACATGATGAGCGGCTGATGATCATCTCC 4502
Db 9643 TCTATGCCAAGAGTACACACAGAGACCTTACAGATGAGCACAAGAAATGATACAGC 9702
QY 4503 ACTATTAATCATCGGATCACTGATCACTGATATGTGCTGTGTGTGCTGCTTGTGCTA 4562
Db 9703 ATTACTACCATCGCATCTGTGTACACATCTTAAGCGGTGACATCACTACCGTGGGGA 9762
QY 4563 TCTGTGTGAGCATGATCATGACAGCTTGTGATGCGCAAGAAAGAGAGTCCCTGA 4622
Db 9763 TGATGATTTGGGTAACTGTTGATGATGATGCTGTGAAAGCGGCGGTGAGTCCCTGA 9822
QY 4623 CGCCATACGCGCTTGACACGAGCGACGATACCGACAGCAATTTAGCGGTTTGTGCTGA 4682
Db 9823 CGCCATACGCGCTTGACACGAGCGCGATATCCAACTTCCTGCTGCTGTGCTGCG 9882
QY 4683 TTGCGCAACCAAGCGTGAACATTTGGAGAAATTTGAAACATCTGTGTTAAACAC 4742
Db 9883 TTAGGTGGCAATGTGAACGTTTACCGAGACATGATTAATCTGTGTGACAGATC 9942
QY 4743 AACGTTTCTGTGGGACAGTTGTGATTCCTCTGCGACCGCTGTATTTCTGTTCCGT 4802
Db 9943 AGCGTTCTTCTGTGGGTCAAGTTGTGATACCTTTGGCGCTTTCATCTTCAAGCCCT 10002
QY 4803 GCTTTTATGCTGATGCTTTTATATGTTGATGAGGCGTCTGCTGAGGAAAGTGAAGC 4862
Db 10003 GCTGCTCTGCTGCTGCTGCTTTTATGTTGATGTTGCGGCGCTTACCTGCGAGAGTGAAGC 10062
QY 4863 CTTTGAACATGACACACTGTGCAAAATGTTTCGGGAGTCCCGTATTAAGCGTGTGCTG 4922
Db 10063 CTTAGGAACATGACACACTGTTTCCAAATGTGCCAGATACCGTATTAAGGACATTTGTTG 10122
QY 4923 AACGCGAGTTTACGCGCACTTAACTGAGATACCGGTGCTCATGGAATTAACAC 4982
Db 10123 AAAAGGCGAGGTATGCGCGCTCAATTTGGAGATCACTGTATGCTCTCGAGGTTTGC 10182
QY 4983 CTTCACTAAAGAGAGTACCTGCAATTCACAGCTATTAATCTTCAACACAG 5042
Db 10183 CTTTCAACCAAGAGATCACTGCAAAATTAACCACTGTGTGCTTCCCAAAA 10242
QY 5043 TTAATGTGCGGCTCTCGAGTGAAGGATCTTCAAGGCGATTAACATGCTCCGCG 5102

```

Db 10243 TCAAATGCTGGGCTCTTGGAAATGTCAGCCGGCCGCTCATGACGATTAATCTGCAAGG 10302
Qy 5103 TTTTGGCGGTGTATACCTTTTCATGTGGGAGAGCGCAATATGCTTCTGTGACGTAGA 5162
Db 10303 TCTTGGAGGGGTCTACCCCTTTATGTGGGAGAGCGCAATGTTTTGGCAAGTAGA 10362
Qy 5163 ACACACAAGTAGAGAGCGGTACGTGAGTTGCTCCAGACTGTACATAAGTACAGCAG 5222
Db 10363 ACAGCCAGATAGAGAGCGGTACGTGAGTTGCTCCAGACTGTACATAAGTACAGCAG 10422
Qy 5223 TCGCACTAAAGTTTACACAGACTGCTCTGAAAGTCGGCTGCTATAGTATACGCAACA 5282
Db 10423 AGGGGATTAAGTGTACACTGCGCGATGAAGTAGAGCTGCTATGTGTAGACGGAACA 10482
Qy 5283 CCACCGGCGCACTGTGATACGTTTGTCAATGGCGTACGCGCGGTTCCTCAGCGGACCTGA 5342
Db 10483 CTACACAGTTTCTTATGATGTGTAGTCGAAACGAGTACACCGAAGACGTTAAACACTTGA 10542
Qy 5343 AGGTCATAGCAGAGCGCGATATCAGCCGCTTTTTCACCTTTGACCATTAAGGTGCTATCA 5402
Db 10543 AAGTCATAGCTGACCAATTTACAGATGCTTTAGCGCTTGTATCATAGTACGTTATCC 10602
Qy 5403 GAAAGGCGCTTTGTTTAACTACAGACTTCCCTGAGTATGAGCTATGAAACAGAGCGT 5462
Db 10603 ATCGGCGCTGTGTATCAACTATGACTTCCGGAATATGAGCGATGAAACAGAGCGT 10662
Qy 5463 TCGGCGATTTCAAGCATCTCTGCTTGTATGCTACAGACATAGTACCGGCACTGACATAC 5522
Db 10663 TTGAGACATTTCAAGCATCTCTGCTTGTATGCTACAGACATAGTACCGGCACTGACATTA 10722
Qy 5523 GGCTGCTGAAGCTTTCTGTCAAGAACATCCAGTCCCTTACACCAAGCACTATCAGAGT 5582
Db 10723 GGCTACTCAAGCTTTCCGCCAAGAACGTGCAATGCTCCGTAACGAGCGCTCATCAGAT 10782
Qy 5583 ATGAATGTGAGAGAACACTCAGAGACGACCCCTGCAAGAAACAGACCATTTGATGTA 5642
Db 10783 TTGAGATGTGAGAAACAACTCAGGCGCCACGACAGAGAAACGCACTTTCCGGTGTGA 10842
Qy 5643 AAATTTGAGTGAAGCTCTGTGAGACGCTTAATCTGTCTTAAGGGCAATCCCTATCTGCA 5702
Db 10843 AGATTTGAGTGAATCCGCTCGAGCGGTGAGCTGTCAATAGGGAAACATTCCTCATTTCTGA 10902
Qy 5703 TTGACATCTCCATGATGCAAGCTTTTGTGAGATCATCAGAAATCAAAATTTTGAAGTTA 5762
Db 10903 TTGACATCTCCGAAAGCTGCTTTTATCAGAGATCAGATGACACACTGCTCTCAACAGTCA 10962
Qy 5763 GCTGCAAGTAGACACTGATTAATTTCTGCAACTTTGTGTCTCTTAACATTAACAGT 5822
Db 10963 AATGTGAAGTCAAGTAGACACTTAATTCAGACAGACTTGGCGGAGATGGCCACCTGCAGT 11022
Qy 5823 ACAAGCTGAGAGGAGGAGATGTCAGATTCCTCCACTCCACGACAGCTGTTTGA 5882
Db 11023 ATGTATCCGAGCGGAGAGGTAAATGCCGTCATTTGCAATTCGAGCAGACAACTCTCC 11082
Qy 5883 AGGAAGCGACCAACATGATGCTGCCGTAGGAGAGATTAACACTATTTTGAACATGCA 5942
Db 11083 AAGAGTGCAGATCATGTCTCTGAGAAAGAGAGCGGTACAGTACACTTTAGCACCGGGA 11142
Qy 5943 GCCCAAGCAAAATTTTATGTTTGTATGCTATGCGGAGAGATGCCAAGTCTGAT 6002
Db 11143 GTCCAGAGGCGAACTTTATGCTATGCTGTGTGGAAAGAAACAACATGACATGAGAAAT 11202
Qy 6003 GTAAACCAACCGGCGGACCACTAATTTGAGAAACCAATTAAGTGCACCAAGATTTCCAG 6062
Db 11203 GTAAACCAACCGGCGGACCACTAATTTGAGAAACCAATTAAGTGCACCAAGATTTCCAG 11262
Qy 6063 CGGCGATTTCCAAAACATCTTGGAACTGGCTGCTTGAACCTGTTTGGGAGACATCATCC 6122
Db 11263 CCGGCGATTTCCAAAACATCATGAGAGTGGCTGTTTGGCCCTTTTGGCGGCGCTCTGTGC 11322
Qy 6123 TCATTTGTGTAGAGATTATGATGTTGTCTGACGCTATGCTTATTAACAACAGCTAGAT 6182
Db 11323 TATTAATATATAGACTTATGATTTTGTCTGAGAGATGATGCTGACACAGAAAGAT 11382

```

```

Qy 6183 GACTGAGCGGCGGACACTGA-CATAGCGGTAAATCTGATGTACTCCGAGAAAGCGTGT 6241
Db 11383 GACGCTAGCGCCCAATGATTCGACAGCAAACTGATGTACTCCGAGAAAGCGTGT 11442
Qy 6242 GCATTAATGCGACGCGCGCT 6261
Db 11443 GCATTAATGATCAAGGCTGT 11462

RESULT 10
AAT30787
ID AAT30787 standard; DNA; 16656 BP.
XX
AC AAT30787;
XX
DT 12-SEP-1996 (first entry)
XX
XX Alphavirus-based eukaryotic layered vector pVDELVIS.
XX
XX Alphavirus; Sindbis virus; vector; gene therapy; vaccine;
XX pVDELVIS; ss.
XX
XX Chimeric Sindbis virus;
OS Chimeric synthetic.
XX
XX Key Location/Qualifiers
FT promoter 1..60
FT /*tag= a
FT /note= "wild-type Sindbis virus sequence
FT capable of initiating transcription"
FT CDS 60..5750
FT /*tag= b
FT /product= non-structural proteins NSP 1-3
FT 5928..7579
FT /*tag= c
FT /product= non-structural protein NSP 4
FT 7579..7602
FT /*tag= d
FT /function= minimal junction region core
FT 11647..11703
FT /*tag= e
FT /function= Sindbis polymerase recognition
FT sequence
XX
XX W09617072-A2.
XX
XX 06-JUN-1996.
XX
XX 30-NOV-1995; 95WO-US15490.
XX
XX 15-MAR-1995; 95US-0405827.
XX 30-NOV-1994; 94US-0348472.
XX 18-JAN-1995; 95US-0376184.
XX
XX (CHIR ) CHIRON VIAGENE INC.
XX
XX Belli BA, Chang SMW, Driver DA, Dubensky TW, Ibanez CE;
XX Jolly DJ, Polo JM;
XX WPI; 1996-27785/28.
XX
XX New recombinant alpha-virus vectors - used to develop prods and
XX methods for use in gene therapy and in the prodn. of vaccines
XX
XX Example 2; Page 186-194; 256pp; English.
XX
XX A eukaryotic layered vector initiation system, pVDELVIS (AAT30787),
XX is based on the Sindbis alphavirus. It is obtd. by inserting
XX the Sindbis clone pVGS6GENrep into vector pCDNA3. The construct
XX can be used in the prodn. of Sindbis virus-based vectors used
XX to develop products and methods for use in gene therapy and in
XX vaccine prodn. pVDELVIS is capable of initiating alphavirus

```

infection; inactivation of the viral junction region (see also
CC AAT30810) will prevent viral transcription of the subgenomic
CC fragment. The efficiency of pVGVIS in generating wild-type
CC Sindbis virus after transfection of BHK cells was approx. 1,000
CC pfu/mg plasmid DNA.

XX Sequence 16656 BP; 4469 A; 4352 C; 4179 G; 3656 T; 0 other;

Query Match 36.6%; Score 2370.8; DB 17; Length 16656;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3694; Conservative 0; Mismatches 1982; Indels 44; Gaps 9;

580 ACTGACGTGTAAGACCGGAGCGGTATATTTCTCATCGGAAACAGCCAGCTACCTTC 639
Db ACTGACTAACCGGGGTAGGTGGTACATATTTTCAGCGGACAGGCCCTGGGCACTTGC 5805
Qy 640 AACCAAAATTCAGTACTCATATTAACAAAGAACTTATTTGGATCGGCGCTTCATG 699
Db 5806 AAAAGAAATGCTGTTGCGAAACGAGCTTACAGAAACGACCTTGGAGCGCAATGTCCTGG 5865
Qy 700 AGAAGTATTAGCGCCCGCGCTCGATCTCGAAAGAGAAATGTTACAGAAAGAACTGC 759
Db 5886 AAGGAATTCATGCCCCGGTCTCGACAGCTCGAAAGAGGAAACACTCAACTCAGTACC 5925
Qy 760 AATTATGCGCTCTGTAAGGAAATAGAACGAGGTATCAATCAGAAAGTAGAAATATGA 819
Db 5926 AGATGATGCCCAACGAGGCAACAAAGTAGTACCAAGTCTGTAAGTAGAAATTCAGA 5985
Qy 820 AAGCAATTAACGCGGAGCGACTATTTCTGGATTGGGCAATATCTATCATAGAAATGA 879
Db 5986 AAGCATATACCACTAGGACGACTACTGTCAGGACTACGACTATTAATCTGCGACAG--- 6042
Qy 880 ATCCGTGAGAGTGTACAGAGTCATTAATCCGTGCAATCTACTACCTGCAAGGTATTA 939
Db 6043 ATCAACGAAATCTATTAAGATACCTATCCGAAACCATTTACTTCAGTACCGGCG 6102
Qy 940 ACAGGTTTACATCTGCAAGAGTGGCGGTTAAACGTGCAACTAGTTATCCAGAGAAAT 999
Db 6103 CGAATCTACCGATCCAGATTCGCTGTAGCTGTCTGTAAACAATATCGATGAGAACT 6162
Qy 1000 ACCCTACGTAGCCAGTTATTTGTATTAACAGATGAATAGATCGGATCTTGACATGTGG 1059
Db 6163 ATCCGACAGTAGCATCTTATCAGATTACTGACAGTACGATGCTTACTTGGATATGGTAG 6222
Qy 1060 ACGGCGCATCGTGTCTAGATACAGCCACTTTTGTCCGGCTAACTGAGAACTACCC 1119
Db 6223 ACGGACAGTGGCTGTCTGATCTGCAACTTCTGCCCCGTAAAGCTTAAAGTTTACC 6282
Qy 1120 CAAGAAACCATAGCTATTTGACAGGAGATTAAGATCAAGCCGCTCCATCGCTATACAGA 1179
Db 6283 CGAAGAAACATGAGATAGAGATAGACCCCGGAAATCCGAGATCGGTTCCATCAGGATGCGA 6342
Qy 1180 ATACATTACAAGATGATTTGGCTGACGCTACTAAAGAAATGCAACGTTACCCAAATGC 1239
Db 6343 ACACCTTCAAAATGTGCTCAATTTGCCGCAACTAAAGAAATGCAACGTCAGAGATGC 6402
Qy 1240 GAGATTTACCTGTCTTAATTTCCGGGCACTTAAATGTTGATGTTTCAAGAAATCGAT 1299
Db 6403 GTGAATCTCCAAACACTGACTCAGGACATTCAAATGTGGAATGCTTGGAAATATGAT 6462
Qy 1300 GCAATGATGAGTACTGAGATACCTTTCGCGATTAACCTATTGGGCTAATCAAGAGAG 1359
Db 6463 GTAAATGAGAGATTTGGAGAGATTCGCTCGGAAACCAATTAAGATTTACACTGAGTTTG 6522
Qy 1360 TTACGCAATATGTGCAAAAGCTGAAAGGCGGAAAGCAGCAGCATTTGTTGGATATCTC 1419
Db 6523 TCACCGCATATGTACTAGACTGAAGAGGCCCTAAGCGCGCACATATTTGCAAGAGAT 6582
Qy 1420 ATTAATCTAAACCGTTGAGAGAGATACCAATGAGATCAATTCGTATGATCTTAAAGAGAG 1479
Db 6583 ATTAATTTGTTCCATTTGCAAGAAAGTCCATAGGATGATTTGCTCATGACATGAAGAAAG 6642
Qy 1480 ATGTCAAGTTACTCTCCGCGACGAACATATACAGAGAGCGGCTAAGGTGACGTTATTC 1539

Db 6643 ACGTAAAGTTATCACACGACGACGAAGAACACAGAGAAAGAACGGAAGTACAAATGATAC 6702
Qy 1540 AGGCTGCAATCCCTTGTCTACCCGTTACTTCTTGGGGATCCATGGGAATTAATGTCCTGA 1599
Db 6703 AAGCGGAGAAACCTCTGGCGATGCTTACTTAATGGGAGATTAACCGGGAATTAATGTCCTGA 6762
Qy 1600 GACTGAATGGGCTGCTTGTCCCAATATTCATACTCTCTTGACATGTCAGCGGAAGATT 1659
Db 6763 GCGTTACGCGCCTCTTGTCTTCCAAATTCATACACGCTTTTGAATATGTCGGGAGATT 6822
Qy 1660 TTGATGCAATTAATGCTGAACATTTTCAACGAGGAGCCAGTATTTGAAACGACATCG 1719
Db 6823 TTGATGCAATATATACAGAAACCTTCAAGCAAGGAGACCCGGTACTGAGAGACGATATCG 6882
Qy 1720 CGTGTGTTATTAAGCGGAAGACGATGTCATTTCCGCTTATGATGATCTTGAAG 1779
Db 6883 CATCATTTGACAAAGCCAAAGCAGACGCTATGCGCTTAACCGGTCATGATCTTGAAG 6942
Qy 1780 ACTTAGGTGTGACCAACCGCTTATAGATTTGATGAGGCGGCTTCCGCAATTCATCAT 1839
Db 6943 ACTGAGGTGTGATCAACCACTACTGACCTTGTATGATGCGCTTTGAGAAATATCAT 7002
Qy 1840 CTGTGACCTTACTTACAGAAACGAGTTTAAATTTGTGCTCATGATGAATCCGATATGT 1899
Db 7003 CCACCATCTACTACGAGGTACTGTTTAAATTTGGGGCGATGATGAATCCGAAATGT 7062
Qy 1900 TCTTAAACGCTGTTTCAACACATGATCATATATCATGATTTCTAGACAGATCTACTG 1959
Db 7063 TCTTCACACTTTTGTCAACACAGTTTGAATGTGTTATCGCCAGACAGATCTAGAAAG 7122
Qy 1960 AAGGTTAACACGCTGACGAGTGGCGGCTCATGAGGAGATTAACATAGTCATGAG 2019
Db 7123 ACGGCTTAAACGCTGACAGATGAGGCTTATTTGGAGCAACATCATATAGAG 7182
Qy 2020 TCGTCTCCGACACTTGTATGAGGAGAGATGCGCACTTGTGTCGATGAATGAAATAA 2079
Db 7183 TAGATCTGACAAAGAAATGCTGAGAGGTGGCCACCTGGCTCAACATGAGGTTAAGA 7242
Qy 2080 TTATGATGACGTTATTTGTATCAAGACACCTTCTGTGGGGATTTATCTGTGGTG 2139
Db 7243 TCATGAGACAGTCACTGCTGAGAGACCACTTACTTGTGCGCGGATTTATCTTGCAAG 7302
Qy 2140 ACCGATTAACGAGCACGCTGACAGATGCGAGACCTCTTAAAGGCTTTTAAGCTTG 2199
Db 7303 ATTGCTTACTTCCACACAGCTGCGCGTGGGAGATCCCTGAAAGGCTTTAAGTTGG 7362
Qy 2200 GAAACCATTTGCACTGATGATACCAAGACTGCGACCGCGCGGCACTGATGATG 2259
Db 7363 GTAAACCGCTCCGAGCGGACGACGAGCAAGACGAAAGAACGCGCTCTGCTAGATG 7422
Qy 2260 AAGCAATGATGAGCAAGAAATGGAATTAAGGACGAGTAAAGGCGCTGAGATCCA 2319
Db 7423 AAACAAAGGCGTGTGTTAGATAGATTAACAGGCACTTTAGCAGTGGCCGAGACACC 7482
Qy 2320 GATACGAAATCATCTGACAGGCTGATCATCAGTCTCTGTCCAGTTAAGCCGAAAGG 2379
Db 7483 GGTATGAGATGACAAATTAACACTGTCTTACCTGCAATGAGAACTTTTGGCCAGAGCA 7542
Qy 2380 TTAAGAACTTAAGAGACTTAAGAGGAGCCCAATCAACCTCTAAGCGGTGACCTTAATAGG 2439
Db 7543 AAGAGCAATTCAGAGCCATCAGAGGGAATTAAGCAATCTTACGCTGTCTTAATAGT 7602
Qy 2440 TGACGTATGACACGACCTTAC-----CCACGGGCAAGATGTTTCCATATCC 2486
Db 7603 CAGCATAGTATCATCTGATTAATCTACTCAACACACACCATGATAGAGATCT 7662
Qy 2487 CTCAGCTGAATTTTCCACCAATTTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 2546
Db 7663 TTAACATGCTGCGCGCGCGCTTCCGCGCGCCCACTGCAATGAGGCGCGGCAAA 7722
Qy 2547 CTCCTAGCGCGCGCTGAGGCGCGCTTCCGCGCGCGCTGCTCAATTCAGATCTTAA 2606

Db	7723	GGAGGACGACGGCCCCGATGCGCTGCCCGGCAACGGGCTGGCTTTCAAATCCAGCAACTGA	7782
Qy	2607	GGAGGTGATAGTCAACTTTGACTTTTCAACAACAGATCACTTAATCCGCGCCAGGTCCA	2665
Db	7783	CCACAGCGCGTAGTGCCCTAGTCAATTGGACAGGCAACTAGACCTTCAACCCCCACGTCAC	7842
Qy	2666	-----CCGCCCAAGAAAGAAAGAGTGTCTCTAAGCCAAAACCTAATAGCTTAAAAAG	2720
Db	7843	GCCCCCACCACCGCCGACGAAGAGCAGGCGCCCAAGCAACCAACGAAAGCCGAAGAAACCA	7902
Qy	2721	AGAAAGCAACGACCAAGAGAGACGAAGC---GGAAGCTTAAACCAAGGGAAGAAAGCAACGTA	2777
Db	7903	AAAGCGAGGAGAAAGAAAGAAAGAACCACTCGCAAAACCCGAAGAAAGACAGCGCA	7962
Qy	2778	TGTGTATGAAAGTTGGAGTCCGACCAAGACATTTCCGATCATG---CTGAACGCGCAAGTGA	2834
Db	7963	TGGCACTTAAGTTGGAGGCGCAGACAGATTGTTGACGTCAAGAAAGAGGAGCGAAGATGTCA	8022
Qy	2835	ATGATATATGCTCGGCTTTCGAGGAGAGGCTGATGAAACCACTCCACTTGAAAGAAAA	2894
Db	8023	TCGGGCAACGCACTGGCCCATGGAAGGAAAGGTAATGAAACCTCTGCACTGTGAAGGAACCA	8082
Qy	2895	TTGATTAATGACCAATTACGGGCGCGGAATTGAAGAGGCTAGACATGACATTTGAGAT	2954
Db	8083	TGCACACCTCTGTGCTATCAAAAGCTCAAAATTTACCAATCGTCAGCAATACACATGGAGT	8142
Qy	2955	ACGGCGACGTTTCCCGAAGACATGAANTCAGACACGCTGCAGTACACACGACAAACACAC	3014
Db	8143	TCGCACAGTTCCAGTCAACATGAAAGATGAGGCACTTCACTTACACAGTGAACACCCCG	8202
Qy	3015	CGGGCTTTCAACACTGGCAACAACGGCGCAGTCCAGTATGAGAAATGGAGATTTACCGTAC	3074
Db	8203	AAGGATTTCTAATACCTGGCAACCAACGAGCGGTGCAGTAAATGAGAGTGAATTTACATCC	8262
Qy	3075	CGAGAGAGTGGGGGGGAAAGGCGCACACGGAAAGCGATCTGGACAAACAGAGGCGAG	3134
Db	8263	CTCGGAGATGAGAGGACAGAGAGACACCGGTCGTCCGATACATGATTAATCTCGGTCCGG	8322
Qy	3135	TTGTGGCTATTGTTCTTAGAGGTGCAATGAGGCGACCGCGTACCGCGCTTTCAGTGTCTCA	3194
Db	8323	TTGTTCGGATATGTCCTCGGTGGCGGTANTGAAGAAACAGAACTGCCCTTTCGGTCTGTCA	8382
Qy	3195	CTTGGAAACCAAGAAAGGGGTGACCATTTAGGATACCCCGAAAGTTCTGAACCGTGT---	3251
Db	8383	CCTGGAATGATGAAGGGAAGACAATTAAGACAACCCCGAAAGGACAGAAAGTGTGTCCG	8442
Qy	3252	-----CACTAGTTTACAGCGCTATGCGGCTTTGCAATGTCAAGTTCCTCCATGCGCAAAAC	3305
Db	8443	CAGCAACACTG3TCAACGGCAATGTGTTGTCTCGGAAATGTAGACTTCCATATGCAACCGCC	8502
Qy	3306	CACCCGTGTGCTAATTCATGACCGCCAGAACGAAACCTCGACGTGCTCGAAGAGAACGTGCG	3365
Db	8503	CGCCCACTGCTATACCCGCGCAACTTTCAGAGCCCTTGACATCTTTGAAGAAAGCTGTGA	8562
Qy	3366	ACAATCCAATTTACGACACGCTGCTGAGAAAGTCTTGAAATGTCCA---TCACGCGCGC	3422
Db	8563	ACCATGAGGCTTACGATACCTGTCTCATATGCAATATTCGCGTGGCGAATCGTCTGGCAAGA	8622
Qy	3423	CCAAACGAAGCATTAACGATGACTTCAACATGACCAAGTCCCTACTCGGGGTTTCGCCCT	3482
Db	8623	GCAAAAGAGGTCGTCTTACGACATTTTACCTTGACACGCGCCCTTACTTTGG3CAATGTCTGT	8682
Qy	3483	ATTGCGAGACCTCAACGCGCGTGTTCACGCCCAATTTAAATTTAGAAACGTGTGGAGCAAT	3542
Db	8683	ACTGCGACCATTACTGAACCGTGTCTTCAAGCCCTGTTTAAGATGAGCAGGTCTTGGACGAAG	8742
Qy	3543	CTGATGATGATCGATTAGAAATCCAGGTCTCGGCACAATTCGGCTTCAATCAGCGACGCA	3602
Db	8743	CGAGAGTAACCACTTACGACATACAGACTTCGCGCCAGTGTGAATACGACCAAGCGGAG	8802
Qy	3603	CTGCGGATGTCAACCAATTCGCTTACATGTCTTTGACCAAGACCATATCAATCAAGAGAG	3662
Db	8803	CAGCAAGCGCAAAACAAAGTACCGCTACATAGTGCCTTAAACAGAGATCACACGTTAAAGAG	8862

QY	3633	ACAGTATGAGAAATAAGCTATACGACATCTTGACACCTCGCCCTGCTTGGCCCAAG	3722
Db	8863	GCACCATGAGATGACATCAAGATATGACCTCAGGACCGTGTGAAAGCTTAGCTACAAAG	8922
QY	3723	GGTACTTCCTGTAGCTCAATGTCTCTCAGGTGACAGTGTAAACCGTCAGTATCAGAGCG	3782
Db	8923	GATCTTCTCTCTCGCAAAATGCCCCCTCAGGGGACAGCGTAAACGGTTAGCATATGTAGTA	8982
QY	3783	GAGCATCTGAGAAATTCATGACCCGTGAGAAAAAGATCAGAGGAAGTTTGTCCGTAGAG	3842
Db	8983	GCAACTCAGCAACGCTCATGTACACTGCGCCGCAAGATAAACCAAAATTCGTGTGGAGCGGG	9042
QY	3843	AGAGTCTTGTGCCACCCGCTCATGAGAAAGCTGTAAAGTGTGCAGTGTATGATCATCT	3902
Db	9043	AAAAATATGATCTCTCTCCGTTACCGGTAAAGAAATTCCTTGACAGTGTACACCCGTC	9102
QY	3903	TGAAGAGACGCTCTGCGGGGTACATAACATGACAGAGCCAGGCCACACGCGTATAGT	3962
Db	9103	TGAAA---ACAACGTGACGGGTACATCATATGACAGAGCCGGGAGCCGACGCTTATACAT	9159
QY	3963	CCTATCTGAGAGAAAGCGTCAGGGCAAGTGTATATAAACCATCTTCTTGCAAGAACGTCA	4022
Db	9160	CCTACCTGGAAGATCATCAGGAGAAAGTTTACGAAAGCCGCACTCTGGAAGAACATTA	9219
QY	4023	CCTACGAATGAAATGTGTGCGCATACAGACACAGGTTATCTGTAGCAGCCGAAACGAATGTA	4082
Db	9220	CGTATGATGTGAATGTGGCGGACATCAAGACCGGAAACGTTTGCACCGCACGAAATCA	9279
QY	4083	ACGCGTGCACTAAAGCAAAACAGTCATTTGCTTCAAGAGCGCAACAAACGAATGGCTCT	4142
Db	9280	CTGGTTGACCGGCATCAAGCAGTGTGCTGCTTAAAGAGGACCAACGAAGTGGGTCT	9339
QY	4143	TCAACTCGCCGGATCTTATTATAGGACACAGACCACTAGTACAGTAAATTTGACATTC	4202
Db	9340	TCAACTACCGGACTTGTATCAGATATAGACACACACAGGCCCAAGGAAATTTGATTTGC	9399
QY	4203	CATTCCGTTGACACCGACAGTCTTGCCTGTTCCGTTAGCTCACAGCCTTACAGTCAAGA	4262
Db	9400	CTTTCAAGTTATCCCGGTGCTGCAATGTGCTCTGTTGCCACAGCCCGCAATTAATAC	9459
QY	4263	AGTGGTTCAAAAGGATACACCTCCACCGACCTGAAATGCGACCAAACTGTGTGCAACGA	4322
Db	9460	ATGGCTTTAAACACATCAGCTCTCCATTAAGATACAGACCACTTACATTTGCTCACACCA	9519
QY	4323	GAAATATGGGGCTCGAGCAGACCGCAACAGCAGAAATGGAATTTACAGGGTCTTACATCCAGA	4382
Db	9520	GGAGCTATAGGGGCAAAACCCGGAAACCAACCATGAAATGATGTGTGAAAGACGGTCAGAA	9579
QY	4383	ATTTTTCTGTGGGGCGAAGAGGCTGGAATCGTATGGGGTAAACCATGAAACCACTCAGAG	4442
Db	9580	ACTTCACCGTGCAGCCGAGATGGCTGTGAATACATATGAGGAAATCATGAGCCAGATGAGGG	9639
QY	4443	TCTGGGGCCAGAGTCCGACACAGGCGACCCCAATGATGGGCCCATGAGATCATCATCC	4502
Db	9640	TCTATGCCCAAGATCAGACACAGAGAACCTTCACGATGTGCCACACAGAAATATATACAGC	9699
QY	4503	ACTATATCATCGGCATTCAGTCTTCACTGTCAATGTGTGTGTGTGTGTGTGTCTTGTCTA	4562
Db	9700	ATTACTACATCGCCATCTGTGTACACCATCTTAGCCGTGCGATCAGCTACCTGTGCGA	9759
QY	4563	TCCATGTAGGACATGCAATCATCAGACAGTTGCGCAAAAGCAAGAAAGACTGCGCTGA	4622
Db	9760	TGATGATTTGGGTACTGTGTGCAATGTGTATATGTGTCTGTAAAGCCGCCGTGATGTCTGA	9819
QY	4623	CGCCATACGCGCTTGCACCGAACCGCAACGATACCAACAGCATTAAGCGGTTTGTGTCTCA	4682
Db	9820	CGCCATACGCGCTTGCACCGAACCGCAACGATACCAACGATTAAGCGGTTTGTGTCTCA	9879
QY	4683	TTTCGGCCACCAACGCTGAAACATTTTGGAGAAATTTTGAACCATCTGTGGTTTAAACAC	4742
Db	9880	TTAGTGTGGCCAAATGCTGAAACGTTTACCCGAGACCAATATGATTTGTGTGTGCAACGTC	9939

QY	4743	AACGGTTCTCTGGGCAAGTGTGACATCTCTCGGAGGCGTTGTAATTCCTGCGCT	4802
Db	9940	AGCGCTTCTTCTGGGTCCAGTTGTGCACACTTTTGGCGGTTCACTGTTCAATCGCT	9999
QY	4803	GCTTTTATGCTGATGCTTTTATTTATGGTGGAGCGCTCTGCTGGGAAAGTATGACG	4862
Db	10000	ACTGCTCTGTGCTGCTGCTCTTTTATGTGTGTCGGGCGCTACCTGGCGAAGTATGAGG	10055
QY	4863	CTTTGAAACATGCGACCACTGTGCGAAATGTTCCGGGGATCCGCTATAAGCGTGTGCG	4922
Db	10060	CCTACGAAACATGCGACCACTGTTTCCAATGTGCGACAGATACCGTATTAAGGCACTGTGG	10119
QY	4923	AACGGGAGGTGTAGCGCGCACTTAACCTGGAGATCAAGGTGCTCATGTGGAAATTAAAC	4982
Db	10120	AAAGGGAGGTGTAGCGCGCTATGCTCAATTTGGAGATCACTGTCAATGTCTCTGGAGTTTTC	10179
QY	4983	CTTCAACTTAACAGAGATGATGATCGTCACTGCAAAATTCACACATGATCTTCTTACCAAG	5042
Db	10180	CTTCCACCAACCAAGAGATCAATTAACCTGCAAAATTCACACATGTGTGTCCTCCCAAAA	10239
QY	5043	TTAAATGCTGGGGTCCCTGAGTGCAGAGCATCTCAAAAGCGGATTCACATGCGCGG	5102
Db	10240	TCAATGCTGGGCTCTTGGAAATGTACGCGGCGCTCATGACAGACTTAACCTGCAGAG	10299
QY	5103	TTTTTGGCGGTGTACCTTTTCAATGGGGAGGCGGACAAATGCTTCTGTAGACATGAGA	5162
Db	10300	TTTTGGAGGGGTGTACCTTTTATGTGGGAGAGGCGCATGTATTTTGGACATGAGA	10355
QY	5163	ACACACACTGATGATGAGGCGTACGTCAAGTTGCGTCCAGACTGCACTATATGATCAGCAG	5222
Db	10360	ACACCCAGATGATGAGGCGTACGTCAATTTGTACAGATTTGCGGCTGTACACACGCGC	10419
QY	5223	TGCGACTTAAAGTTTACACACAGCTGCTGTGAAGTGGCGCTGCGTATAGTATACGCGACA	5282
Db	10420	AGGCGATTTAAGTGTACACACTGCGCGCATGAAGATGAGACTGCTATATGTATCGGAAACA	10479
QY	5283	CCACCGGCGACCTGGATACGTTTGTCAATGGCGTCAAGCGGATTCCTACGGGAACTGA	5342
Db	10480	CTACCACTTTTCTAAGATGTATCGTGAACGAGTCAACACAGAAAGCTCTTAAAGACTTGA	10539
QY	5343	AGGTCAATGACAGGGCGATATCAAGCGCTTTTTCACCTTTGACCAATTAAGTGTGATGA	5402
Db	10540	AAAGTCAATGCTGAGACCAATTTTCAAGATGTTTACGCAATTCATTAAGTGTGTTATCC	10599
QY	5403	GAAAGGGGCTTGTTTTCAACTACGACTTCCCTGATGATGAGGCTATGAAACAGAGCGT	5462
Db	10600	ATCGCGGCGTGTGTATCAACTATGACTTCCCGGAATATGAGGAGTAAACAGAGCGCT	10655
QY	5463	TGCGCGATTTTCAAGCATCTCTCGCTTATGCTACAGACATATGATGCCCGCATGACATAC	5522
Db	10660	TGCGAGACATTTCAAGCATCTCTCTTGTACTGACAGAGATCTCATGCGCAGACAGACATTA	10719
QY	5523	GGCGTGTGAAGCCCTTGTCTAAGAAACATTCACGCTCCCTTACACCCAGACAGATTCAGGGT	5582
Db	10720	GGTACTCTAAGCCCTTCCGCAAGACGTGATGTCCTGTACAGCAGCGCGCATACGAT	10779
QY	5583	ATGAATGTGGAAGAACATCTCAGAGCAACCCCTGACAGAAACAGACCAATTTGATGTA	5642
Db	10780	TTGAGATGTGGAAAAAACAATCAGGCGCGCCATCGACAGAAACCGACCTTTGGGGTGA	10839
QY	5643	AAATTTGAATGTGAGCCCTCTGTGAGCGCTTAACTGTGTGTTAAGGGCAATCCTATTCGA	5702
Db	10840	AGATTTGACATTAATCCCTCTCCGAGCGGTGACGTTCATCTCGGGAACATTTCCATTTCTA	10899
QY	5703	TTGACATCCCGATGACAGCTTTTGTGTGATCATCAGATACCAACAATTTTGAAGTTA	5762
Db	10900	TTGACATCCCGAAGCGCTGCTTTATTCAGACATCAGTGCACCACTGTCTCAACAGTA	10959
QY	5763	GCTGCACATGACAGACTGATTTATTTCTGACAGCTTTGGTGTGTTCTTAACTTACAT	5822
Db	10960	AATGTGAAGTCAATGATGTGACTTATTAACAGACACTTCCGCGGGAGATGGCCACCTGTCACT	11019
QY	5823	ACAAAGCTGACAGGAGGACATTTCTCACTTCCATCCACGACAGCTGTTTGA	5882

Db	11020	ATGATATCCGACCGGAAAGTCAATGCCCGTACATTGGCATTTGAGCACAAGCAACTCTCC	11079		
Qy	5883	AGGAAGGCACACACATGTAAGTACCCGTAGGCGCAGCATTAACACTACATTTTACACATCGA	5942		
Db	11080	AAGAGTGCACAGTACATGTCCTGTGGAGAAAGACGGGTGACAGTACACTTTAGCACCGCGA	11139		
Qy	5943	GCCCCACAAGCAAAATTTTATAGTTTGGCTATGCGGCAAGAAAGTCCACCTGCAATGCTGAAT	6002		
Db	11140	GTCACACAGCGGAACCTTATGTGTGCTGTGTGGAAAGAAACATGCAATGCAGAAAT	11189		
Qy	6003	GTAACACACCGCGCGACCATATTAATTTGGAGAACCATATAAAGTCAGCAAGAAATTCAGG	6062		
Db	11200	GTAACACACACAGCTGACCATATGTGAGACACCCCGACAAATAATGCCAAGAAATTTCAAG	11255		
Qy	6063	CGGACATTTTCCAAACATCTTGGACCTGGCTGCTTGCACATGTTTGGGGAGCATATCCC	6122		
Db	11260	CCGCGCATCTCAAAAACATCATGAGTAGTGGCTGTGCTTTCCTTTTCCGGCGGCGCTCGTCC	11319		
Qy	6123	TCATTTGTTAGAGACTTATATAGTGTGGTGTGACGCTCTATGCTTATTAACACAGCTGAAT	6182		
Db	11320	TATTAATTTATAGACTTATATGATTTTGTCTTGACAGCATGATGCTGACTACACAGAAAT	11379		
Qy	6183	GACTGACCGGGGACACTGA-CATAGGGGTAAATCTGATGTCCTTCCGGAAGCGTGGT	6241		
Db	11380	GACCGCTACGCCCCCAATGATCCGACACAGCAAAACTGATGTACTTCCGAGGAATGATGT	11439		
Qy	6242	GCATTAATGCCACCGCGCGCT	6261		
Db	11440	GCATTAATGCATCAGGCTGCT	11459		
RESULT 11					
AAAT0577					
ID	AAAT0577	standard; cDNA; 11717 BP.			
AC	AAAT0577;				
DT	06-DEC-2000	(first entry)			
XX					
DE	Sindbis-like virus strain YN87448 complete genome.				
XX					
KW	Genome; Sindbis-like virus strain YN87448; primer; RT-PCR; vaccine;				
XX	epidemic; Sindbis encephalitis; evolution; epidemiology; ds.				
OS	Sindbis-like virus strain YN87448.				
XX					
PN	CN1252445-A.				
PD	10-MAY-2000.				
XX					
PF	27-OCT-1998; 98CN-0120694.				
PR	27-OCT-1998; 98CN-0120694.				
XX					
PA	(VIRO-) INST VIROLOGY CHINESE ACAD PREVENTIVE ME.				
XX					
PI	Liáng G, Zhou G, Li L;				
XX					
DR	WPI; 2000-443226/39.				
XX					
PT	Whole genome sequence of YN87448 virus strain and its cloning method -				
PS	Claim 1; Page 2-9; 24pp; Chinese.				
XX					
CC	This sequence represents the complete genome of the Sindbis-like virus				
CC	strain YN87448. The genome was cloned as 15 fragments using the PCR				
CC	primers AAAT0578-A700603 into the plasmid pGEM-T. The invention relates				
CC	to the isolation and method of cloning the complete genome for the				
CC	Sindbis-like virus strain YN87448 by a RT-PCR process. The YN87448				
CC	strain virus appears to be the optimal candidate for a vaccine to				
CC	prevent epidemics of Sindbis encephalitis. The sequence of this				
CC	strain's genome shows the difference between this viral strain and				

CC other epidemic Sindbis virus strains at the molecular level and is
CC useful for understanding the source, evolution and molecular
CC epidemiology of Sindbis viruses.

XX Sequence 11717 BP; 3300 A; 3109 C; 2924 G; 2384 T; 0 other;

Query Match 36.6%; Score 2370.6; DB 21; Length 11717;
Best Local Similarity 64.3%; Pred. No. 0;
Matches 3737; Conservative 0; Mismatches 1999; Indels 77; Gaps 9;

```
QY 580 ACTGACGATGAAACCGGAGCGTATATTTTCATGGAACAGGCCAAGCTACCTTC 639
DB 5758 ACTGACTACACCGGGGTGAGTGGGTACATATTTTCAGAGGACAGGCCCTGGGCACTTTCG 5817
QY 640 AACGAATCAGTACGTCATGTAATGTAACAGAACCTATATTGGATCGGGCGCTCATG 699
DB 5818 AAAAGAACTCGTCTGAGAACCAAGCTTACAGAACCGACTTGAGGCGCAATTTCTCG 5877
QY 700 AGAAGTATACGCCCCCGCTCGATCTCGAAGAGAGAAATGTTCAGAGAACTGC 759
DB 5878 AAGAAATCTAGCCCCCGGTGCTGACACGTGAAAGAGAACACTCAATCAGTACC 5937
QY 760 AATTATCGCTCTGAGAAATAGAAACAGATATCAATCAGAAAAGTAGAAATATGA 819
DB 5938 AGATGATGCCACCGAAGCCAAACAAAGCAGTACAGTCTGAAAAGTAGAAAACAGA 5997
QY 820 AAGCAATTAACGGGACGACCTATTTCTGATTTGGGACACATCTATCATCAGAACTGA 879
DB 5998 AAGCAATTAACCGACGACGATGCTTTTCAGGGCTACAGCTGTAAATCTGCGCACAG-- 6054
QY 880 ATCTGTGAGTGTATACAGATCAATTAATCTGTACCAATCTACTCTGTCACGCTAATTA 939
DB 6055 ATCAGCCAGATGCTATTAAGTACACCTACCGAAACCATGCTATTCAGACGTTACAG 6114
QY 940 ACAGGTTAATCTGACAGAGTGCCTGTTAAACGTCACTTAATTCAGAGAAAT 999
DB 6115 CTAACTCTGACCCCAAGTTTCTGTAGCTGTTTGAACAATCTGTGATGAGAAAT 6174
QY 1000 ACCCTACAGTACAGATTAATGTAATCAAGTGAATCGATCTTGAACATGATG 1059
DB 6175 ACCGACGATGACATCTTATCAGATCACGACGAGTACGATGCTTACTGATATG 6234
QY 1060 ACGGCGATCGTGTCTGATATACAGCACTTTTTCGCGCTAACTGAGAACTACC 1119
DB 6235 ACGGACAGTGCCTTGTGCTGATATCTGCAATTTTTCGCCCGCAAGTTAGAAATACC 6294
QY 1120 CAAGAAGCATGCTATTTTTCAGCCAGATTAAGATACGCGTCCATGCTTATACGA 1179
DB 6295 CGAAAAGACACGAGTATAGACCCCAACATCCGACGTGCTTCATCAGCGATGAGA 6354
QY 1180 ATACATTACAAATGTAATGCTGCTAGCTACTAAAGGAATGGAAGGTTACCAAAATGC 1239
DB 6355 ACAGTTGCAAAACGTGCTACTTCCCGACTTAAAGAACTGCAACGTCACAAATGC 6414
QY 1240 GAGATTAACCTGTCTGATTCGCGGCAATTAATGTTGATGTTTCAAGAAATACGAT 1299
DB 6415 GTGAATCGCAACACTGCACTCAGCAGATTCACAGTTGATGCTTTCGAAAATATGAT 6474
QY 1300 GCAATGATGATCTGGGATACCTTTTCGCGATTAACCTATTCGCTTAACTACAGAAACG 1359
DB 6475 GCAATGACGAGTATTTGGAGGAGTTTTCGCAAAAGCCAAATTAGATCACTCTAGATTG 6534
QY 1360 TTAGCAATATGTCACAAAGCTGAAAGGCGGAAAGCAGCATGTTTTCGGAATCTC 1419
DB 6535 TTACCGCATACGTGCGCAGACTGAAGGCGCTTAAGCGCGCGCATGTTTCGCAAAAGCG 6594
QY 1420 ATATCTTAAACCGTTCAGAGAGATTAACATGATGATCAATTCGTATGATCTTAAAGAG 1479
DB 6595 ATATATTGCTCCATCTGAGAAAGTGGCTATGATGATGATGCTATGACATGAAAGAG 6554
QY 1480 ATGTCAAAATTAATCTCCGCGACGAAACATACAGAGAGCGGCTTAAGGTGAGTTATTC 1539
DB 6555 ACGTGAAGTTTACACCTGCGACGAAACACACAGAAAGAAAGCCGAAAGTACAAATGATAC 6714
```

```
QY 1540 AGGTCGATATCCCTTGTACCGCTTACCTTTGCGGATATCATCGGAATTAATCCGTA 1599
DB 6715 AAGCCGACAAACCTCTGGCCACCGCTTACTTATGCGGATTCACCGGATTAATGTCGCA 6774
QY 1600 GACTGATGCGGTGCTTTCGCAAAATATCCATCTCTTTCGACATGTCAGCGGAAT 1659
DB 6775 GCGTACAGCGTTTGTCTACCAACATTCACACGCTCTTGAACATGTCGCGGAGACT 6834
QY 1660 TTGATGCGATTAATGCTGAACATTTCCACACGCGCAACCAATTTGGAACGCACTCG 1719
DB 6835 TTGATGCAATCAATAGCAAGAACCTTCAAGAAAGTGAACCGTATCTGGAACGATATCG 6894
QY 1720 CGTGTGTTGTAAGAGGAGAGAGAGAGCTATTCCTATTTGCGGCTGATGATCTTGAAG 1779
DB 6895 CCTGCTTGCACAAAGCCAGAGAGAGCGCTATGCGTTAACCGGCTGATATCTTGGAG 6954
QY 1780 ACTTAAAGTTCGACCAACCGCTCTTGAATTTGATAGAGCGGCGTTGCGCAATATCAT 1839
DB 6955 ACTGGGTGTCAGCAACCACTACTGACTTGAATGAGTGCCTTTGAGAAATATCAT 7014
QY 1840 CTGTGACCTTACTACAGAAACGAGTTTAAATTTGCTGATGATGAATCCGCTATCT 1899
DB 7015 CCAACCATCTGCGCACGCGGTACCGCTTCAAAATTCGGGCGATGATGAATCCGGAATGT 7074
QY 1900 TCTTAAAGCTGTTTGTCAACACACTAGTCAATATCATGATTTGCTAGCAGAGTACTACGTG 1959
DB 7075 TCTTCAAGCTCTTTGTCAACACAGTTTGAATGCTGTTATGCGCAGCAGATTTGAGAG 7134
QY 1960 AAGCGTTAACACGTCAGCGTGCAGCGCTCTATTCGGCAGATTAACATATGATGATG 2019
DB 7135 AAGCGCTTAAACGTCCAATATGTCAGCATTTATTCGGCAGATTAACATATACACGAG 7194
QY 2020 TCGTCTCCGACACCTTATATGCGGAGATGCGCCACTTGTGGTGAACATGGAAGTAA 2079
DB 7195 TAGATCTGCAAAAGAAATGCTGAGAGGTGTCACACTGCTCAACCTGAGAGTTAGA 7254
QY 2080 TTAATGATGAGTATTTGATATCAAGCACTTACTGTTGGGAGATTTATCTGTG 2139
DB 7255 TCATTGACGATGATGAGGAGAGACCACTTACTTCTGCGGTGATTAATCTTGAAG 7314
QY 2140 ACCAGATTAACGACACAGCTGACAGTGCAGACCTCTTAAAGGCTTTTAAAGCTTG 2199
DB 7315 ATTGCTTACTCTCAACGCTGTGCGGTGCGAGACCTTGAAGGCTTTTAAAGTTGG 7274
QY 2200 GAAACCATTTGCGATGATGATACCAACATCTGCAACCGCCCGCGGACCTGATGATG 2259
DB 7375 GTAAACCGCTCCAGCGCAGACGAGCAGACAGAAAGCAGAAAGCGGCTCTGCTATAG 7434
QY 2260 AAGCAATGCGATGAACAGAAATGGAATTAACGACGATGATGAGGCGGTAGATCCA 2319
DB 7435 AAACCAAGGCGTGTGTTAGATGATTAACAGACACCTTATGACAGTGGCGGTGCACTC 7494
QY 2320 GATACGATCATATCGCAGCGCTGATCATAGTCTGTCTCCAGCTTAACCGAAAGCG 2379
DB 7495 GGTATGAGTACAAACATCAACACTGTCTGCTGAGTATGAAGAACTTTGCCAGAGCA 7554
QY 2380 TTAAGAACTTCAAGAGATTAAGAGGAGGCCATACCTCTACGCTGACCTTAAATAG 2439
DB 7555 AAAGAGCATTTCAAGCATCAGAGGAGAAATTAAGGCACTCTACGCGGTCTTAAATAGT 7614
QY 2440 TGACGTAGTAGA-----CAGCACTTACCAACCGGAGCAATGTTTCCA 2482
DB 7615 CAGCATATGATATTTATCTGATTAATACACACACACACATGATTAAGAGATTTCT 7674
QY 2483 TACCTTACGCTGAATTTTCACAGATTTTACCTTAACCAATTCGATGCTTACGAGATCCA 2542
DB 7675 TTAACATGCTCGGCGCGCGCTTCCAGCCCGCCCATGECATGATGAGGCGCGGAGAA 7734
QY 2543 AACCTTCTAGGCGCGCTGAGAGCGCTTTCGCGCCCGCTGCTGCTCAATTCGAAGAT 2602
DB 7735 GAGGACGAGGCGCGCTGATCTGCGCAGCAATGGCGTTCCTCAATTCAGACACTGA 7794
```

QY 2603 CTTAGAGGTGATAGTCAACTTGACTTTCAAAC--GATCAGCTATCCGCCGAG 2660
Db 7795 CCACAGCGGTGAGTCCCTAGTCAATGAGCAGCAACTAGACTTAAACCCACGCCAC 7854
QY 2661 GTCCACCGCCAAAGAAAGAAAGTGTCTTAAAGCCTTAACTTAAAG 2720
Db 7855 GCCCCCGCCCGCAGAAAGAGGCGCCAAAGCAACCGAAGCGAAGAAACCA 7914
QY 2721 AGAAGCAGCAAGCCAAAGAGAGCAAGCG--AAGCTTAAACAGAGAAAGCAAGTA 2777
Db 7915 AAACACAGAGAAAGAAAGAAAGCACTGCAAAACCCAAACCCGAAAGACAGCGTA 7974
QY 2778 TGTGTATGAAGTTGAGTCCGACAAAGATT--TCCATTCATGCTGAACGCCAAGTGA 2834
Db 7975 TGGCATTAAGTTGAGGCCGACAGACTGTTGAGTAAATAAGAGCGAGATGTCA 8034
QY 2835 ATGATATGCTGCGCTTGTCCGAGAGAGGCTGATGAACCACTCCAGTTGAAGAAAA 2894
Db 8035 TCGGCGACGCACTGGCCATGGAAGAAAGGTAAATGAACCTCCACGTAAGAAAGCA 8094
QY 2895 TTGATATGAGCAATTAGCGCGCCGTAATTGAAGAGCTAGCATGTAAGCACTTGGAGT 2954
Db 8095 TTGACACCGCTGTCTATTAAGCTCAATATTCAGAGTCTGACATTCAGCATAGAGT 8154
QY 2955 ACGGCGAGTTCCCGAAGCATGAATGAGACACGCTGAGTACACAGCAAGCAACAC 3014
Db 8155 TCGCACTATGCGGTGCAACATGAGAGAGCGTTTACCTTACACAGTGAACCTCG 8214
QY 3015 CGGCTTTTACACTGGCAACACGCGCGAGTCCAGTATGAGATGAGATTTACCGTAC 3074
Db 8215 AAGGTTTCTACAACTGGCAACACGAGCGGTGCAATATGAGAGCGATTTACATCC 8274
QY 3075 CGAGAGGATGGGGCGGAAAGCGCAGCGGAAGCCATCTCGACAAACAGAGCAAG 3134
Db 8275 CCGCGGATGAGAGGACAGAGAGACAGTGTGTCGATTAAGATTAATCAGAGCGAG 8334
QY 3135 TTGTGCTATTGTTCTAGAGAGTGCATAATGAGGACGCGTACGCGCTTTCAGTGTCA 3194
Db 8335 TTGTGCGATAGTCTCGAGAGGGGTGATGAGGAAACAAAGACCCCTTTCGTCGTA 8394
QY 3195 CTTGGAACCAAGAAAGGGGTACCAATTAGGATATCCCCGGAAGTTCTGAACCTGT 3251
Db 8395 CCGTAATAGCAAGAAAGGAAAGCAATCAAGCAACCCCGAAGGACAGAGAGTGTCTG 8454
QY 3252 -----CACTAGTTACAGCGCTATGCGCTTTCGAAATGTCACTTCCCATGCCAAG 3305
Db 8455 CTGACCACTGTGTCCGCGCATGTCTTGTCTGAAACGTGAGCTTCCATGCAATGCC 8514
QY 3306 CACCGTGTCTATTCACTGACGCGAAGCAACACTGACGTGCTCGAAGAGAGCTCG 3365
Db 8515 CGCCCACTGTCAACCGCGAACCATCAGAGCTCTGACATCTCTCGAAGAGACGTGA 8574
QY 3366 ACAATCCAATTTACAGACGCTGTGAGAAAGCTTGAATGTCAATC-----ACGC 3419
Db 8575 ACACAGAGGCTTACGACACCGCTGCTCAAGCATTTGGGGTGGGATCGTCCGCAAG 8634
QY 3420 GCGCCAAACGAAGCTTACCGATGACTTCACTGACAGTCCCTTACCTGGGCTTCTGC 3479
Db 8635 GTAAAAAAGAGCTCACTGACGCTTACCTTACCAAGCCCGTATCTTGGGCAATCT 8694
QY 3480 CGTATTGACAGACTCAAGCGCGTGTTCAGCCCAATTAATAATGAGAACGTTGGAG 3539
Db 8695 CGTACTGTACCATCTGAACCGTCTTAACTTAAATGAGTCAAGAGGTCTGGAGTGG 8754
QY 3540 AATCTGATGATGATGATGATTAAGATTCAGGCTTCGCAATTCGCTTCAATTCAGGAG 3599
Db 8755 AAGCGAGAGCAACACCATACGATACAGACTTCGCGCAAGTTTGGATACGACCAAGG 8814
QY 3600 GCATCGCGAGTGTCAAAATTCGTTACATGTCTTTGACCAAGACATGACATCAAGG 3659
Db 8815 GAGCAGCAAGCTCAAAATTAAGTACCGCTTACATGACCTCGAGCAGATCATATCTGCAAG 8874
QY 3660 AAGACATGATGAGAAATAGCTATCAGACATCTGAGACCTGCGGTGTTGGCACA 3719

Db 8875 AAGGACCATGATGATGATCAAGATCAGACACTGAGACCGGTGGAAGGCTTACTACA 8934
QY 3720 AAGGTACTTCTGTACTCAATGTCTCTCGAGGTGACAGTGTAAACGCTAGTATCAGA 3779
Db 8935 AAGGATCTTCTCTCTCCGAAAGTGTCTTCCAGGGGAGCAGGTATCGGTATGACATGCA 8994
QY 3780 GCGGAGCATCTGAGATTAATTCATGACCGTGAAGAAATATCATGAGAGAGTTGTCCGTA 3839
Db 8995 GTAGCACTCAGCAACGCTATGACATATGAGCCCGCAATTAACCAAAATTCGTGGAC 9054
QY 3840 GAGAGAGTACTTGTTCACCGCTCCATGGAAGCTGTAAAGTCCACGTTTACGATC 3899
Db 9055 GGAATAAATATGACCTTACTCCCGTTTCAAGAGATTCCTTGCAAGTGTACGACC 9114
QY 3900 ACTTGAAGAGACGCTTCGCGGTATCAATACATGCAACAGCCAGGCCACACGCTATA 3959
Db 9115 GTCTGAAAGAAACAAACCCCGCTACATCACTATGCAAGGCGCGGACCGCCTATA 9174
QY 3960 AGTCTATCTGAGAGAGCGTCAAGCGAGTACATTAACCACTTCTGGCAAGAACG 4019
Db 9175 CATCTTATCTGAAGAAATCATAGGAAAGTTTACGGAAGCCACATTCGGAAAGAAC 9234
QY 4020 TCACCTAAGAAATGATAGTGTGCGATACAGACAGTATCGTGAACAGCGAAGAA 4079
Db 9235 TTAGGTATGAGTGAAGTGGCGGATTAAGAACCGGAACGTTTACGACCGGTACGAAA 9294
QY 4080 TGAACGCTGCACTTAAGCAAAACATGTCAATTGCTTACAAAGAGCAACCAAAATGGG 4139
Db 9295 TCACGCGCTGACCGCCATCAAGAGATGCTGCTGCTTAAAGCCAAACGAAGTGGG 9354
QY 4140 TCTTCAATCGCGGATTTTATGAGCAACAGACCACTCAGTGTGAAGTAAATTGCA 4199
Db 9355 TCTTCACTGCGGACTCTATCAGACAGCGCCAGACACAGCCCAAGGAAATGTGCA 4259
QY 4200 TTCAATTCGCTTGCACACGACAGTGTGCGGTTCCGTTAGCTTACACAGCTTACGTA 4259
Db 9415 TGCCTTTAACTGATCCCGAGTACCTGATAGTGTCCGTTGCGCACGCGCGAAGCTAG 9474
QY 4260 CGAAGTGTCAAAAGGCTACACCTTCACTGATGCAATGCAACCAATTCGTGACAA 4319
Db 9475 TACACGCTTTAAACATCACTGAGCTTCAATTGACACAGACCATCTGACATTCGACCA 9534
QY 4320 CGAAGAAATTTGGGGGTGAGACAGACAGACAGAGATTAACAGGCTTACATCA 4379
Db 9535 CGAGAGCTAGGGGCAAAACCGGAACCAACATCAAGATGATCTCGAAAGAGGTTA 9594
QY 4380 GAAATTTTCTGTGGGCGAAGAGGCTGAGATAGTATGGGTTAACATGAAACAGTCA 4439
Db 9595 GAACTTCAACGCTGACCGGATGAGCTGGAATATATATGGGCAATACAGAACAGTAA 9654
QY 4440 GAGTGTGGCCGAGAGTGGCACCAAGGCAACCAATGATGGCGGATGATCATCA 4499
Db 9655 GGGTCTATGACCAAGAGTCTGACCAAGGAGACCTTCAAGATGGGCAACAGAAATAGTAC 9714
QY 4500 TCCACTTATATCATGAGATTCAGTCTTACATGTCATATGTGTGTGTGTGTGTGTGT 4559
Db 9715 AGCATTTACTATCATGCGCATCTGTGTACACATCTTAAGCGGTGATCAGTGTGTGTG 9774
QY 4560 CTATCTGTGAGGCACTGATCATGACAGCTTGCATGCGCAAGCAAGAGAGAGTGC 4619
Db 9775 CGATGATATGAGGATATGTTGAGATTAATGTGCTGTAAAGCGCGGTGAGTGC 9834
QY 4620 TGAAGCATATGAGCGCTTGCACCGAAGCGAAGGATACCAAGCATTAAGCGGTTTGTCT 4679
Db 9835 TGAAGCATATGAGCGCTTGCACCGAAGCGAAGGATTAAGCATTAAGCGGTTTGTCT 9894
QY 4680 GCATTCGCGCAACCAAGCGTAAACATTTGAGAACTTTGAAACATGTGTGTGTAA 4739
Db 9895 GTGTAGGTGTGCTATGCTAAACATTCACCGAGACCATAGTTACTTATGTGTGAA 9954
QY 4740 ACCAACGTTTCTGTGGGACAGTTGTGATCTCTGTGGACGCGCTGTATTCGTGTCC 4799

Dh 9955 GCCACCGCTTCTTCGGGTCCAGCTGTGTATACCTCTGCGCGTGTATCTGCTTAATGC 10014
Qy 4800 GCTGCTTTTCAATGCTGATGCTTTTATTTGTTGAGGCGTGTGCTGGGGAAGTAG 4859
Dh 10015 GCTGTGTATGCTGCTGCTCTTTTATGTTGTGCGGCGCTACCTGGCGAAGTAG 10074
Qy 4860 ACGCTTTCGAACATGCGACCACTGTGCCAAATGTTCCGGGAGATCCCGTATAGGCGTTGG 4919
Dh 10075 ACGCTTCGAACATGCGACCACTGTGCCAAATGTTCCGGGAGATCCCGTATAGGCGTTGG 10134
Qy 4920 TCGAACCGGAGTTTATGCGCGCACTTAACCTGAGATCACGGTGCTCATGGAATTA 4979
Dh 10135 TTGAAAGGGGCGGGGTAGCGCCCGCTCAGTTGGAGATTACTGTATGTCTCTGTAAGTTT 10194
Qy 4880 CACCTTCACTAACAGGAGATGACCTGCAATTCACACAGTCACTTCTTCCACAC 5039
Dh 10195 TGCCTTCCACCAACCAAGATACATCACTGCAAAATTCACACGTGTGCTCCCTCCCTTA 10254
Qy 5040 AAGTTAATGCTGCGGGTCCCTCGAGTGCAGAGCATCCTCAAAAGCGGATTACACATGCC 5099
Dh 10255 AAGTCAATGCTGCGGGTCCCTCGAGTGCAGAGCATCCTCAAAAGCGGATTACACATGCC 10314
Qy 5100 GCGTTTTTGGCGGTGTGATCCCTTTCATGTGGGAGCGCAAAATGCTTCTGTGACATG 5159
Dh 10315 AGGTCTTTGGAGGGGTGTACCCCTTCATGTGGGAGCGCAAAATGTTTGGCAGATG 10374
Qy 5160 AGAACACACACATGATGAGGCGGTACGTCGAGTTCCGCTCCAGATGCACTTAATGATACG 5219
Dh 10375 AGAACACACATGATGAGGCGGTACGTCGAGTTCCGCTCCAGATGCACTTAATGATACG 10434
Qy 5220 CAGTGCACATAAAAGTTTACACACAGCTGCTGAAAGTGCAGCTGATATGATATACGCA 5279
Dh 10435 CGCAGGGATTAAGGTGATCTGCGCGATGAAAGTATGAGCTGCTATATGATGAGGA 10494
Qy 5280 ACACACCGCGCACCTGGATACGTTTGTCAAATGCGTCAACCCAGTCTTCCACGGAGC 5339
Dh 10495 ACACATACATTTCTTCTAGATGTGTACGTGAACGAGTACACCGAAGCTCTTAAGACC 10554
Qy 5340 TGAAGTTCATGACGAGGCGGATACAGCGGCTTTTTCACCTTTCACCATTAAGTGTGCA 5399
Dh 10555 TGAAGTTCATGACGAGGCGGATACAGCGGCTTTTTCACCTTTCACCATTAAGTGTGCA 10614
Qy 5400 TCAGAAAGGGGCTGTTTACAACTACAGACTTCCCTGAGTATGAGCTATGAACAGAGAG 5459
Dh 10615 TCCATCGGCGCTGTGTACAACTATGACTTCCGGAATAGAGGCAATGAACAGAGAG 10674
Qy 5460 CGTTGGGATATTCAGCATCTCTGCTTGAATGCTACAGACATAGTAGCCCGCATGACA 5519
Dh 10675 CGTTGGGATATTCAGCATCTCTGCTTGAATGCTACAGACATAGTAGCCCGCATGACA 10734
Qy 5520 TACGGCTGCTGAAGCTTCTTCTCAAGAACATCCAGTCCCTTACACCCAAAGCATATCAG 5579
Dh 10735 TTAAGCTATCAGCTTCTTCCGCAAGAACGTGATGCTCCGTACACGAGCGCATCTG 10794
Qy 5580 GGTATGAATGTGGAAGAACAACTCAGAGACAGCCCTGCAAGAAACAGACATTTGGAT 5639
Dh 10795 GATTCGAGATGTGGAAGAACAACTCAGAGACAGCCCTGCAAGAAACAGACATTTGGAT 10854
Qy 5640 GTAAATTTGAATGAGAGCTCTTGCAGAGCTTAACTGTGCTTAAGGAGCAATCCCTATCT 5699
Dh 10855 GCAAGATTTGAGATATCCGCTTCAAGGCGGTGAGCTGCTCAAGGGAACATTTCCCATCT 10914
Qy 5700 CGATTGACATCCCGATGACACTTTTGTGAGATCATCGAGATCAACCAATTTTAAAGAG 5759
Dh 10915 CTATTGACATCCCGATGACACTTTTGTGAGATCATCGAGATCAACCAATTTTAAAGAG 10974
Qy 5760 TTACTGACAGTAGACAGACTGATTTATTTCTGAGACTTTGTGTTCTTAACTATAC 5819
Dh 10975 TCAATGTGATGTAGTAGAGTCACTTATTCAGGGAATTTGGGCGGATGTGCTACCTGC 11034
Qy 5820 AGTACAAAGTACAGGAGGAGCAATGTCCAGTTCACTCCCATCCAGACAGCTGTTT 5879
Dh 11035 AGTATGTATCGAGCGGAGGAGCAATGTCCCTGTATTCATTCGATTCGAGACAGCAACCC 11094

Qy 5880 TGAAGAGGACACACATGATGATGCTCCGTAGGACAGATACATCATTTTACACAT 5939
Dh 11095 TCCAGAGTGTGACAGTTTATGCTCTGTGAGAAAGAGGCGTGTACATCTTACGACCG 11154
Qy 5940 CGAGCCCAAGCAAAATTTATGATTTGCTATGCGGCAAGAAATGCTCACTGCAATGCTG 5999
Dh 11155 CGAGCCCAAGCAAAATTTATGATTTGCTGTGTGTGTAAGAAACAAATGCAATGACAG 11214
Qy 6000 AATGTAAACACCGGCGGACACATTAATGGAACACACATTAAGTGCACCAAGATTC 6059
Dh 11215 AATGTAAACACCGGCGGACACATTAATGGAACACACATTAAGTGCACCAAGATTC 11274
Qy 6060 AGGCGGAGTTTCCAAACATCTTGAACGTGCTGCTTGAACATGTTTGGGGAGCATCAT 6119
Dh 11275 AAGCGGCATCTCAAAACATCTTGAACGTGCTGCTTGAACATGTTTGGGGAGCATCAT 11334
Qy 6120 CCTCATTTGTTAGAGCTTATAGTGTGTGCTGACGCTTATGCTTATTAACACAGCTA 6179
Dh 11335 CGCTATTATTTATGAGCTTATGATTTTCTTGCAGACATGATGCTGACTAGACAGAA 11394
Qy 6180 GATGACTGAGCGGGAACCTGA-CATAGCGTAAATCGATGACTTCCGAGGAAGCT 6238
Dh 11395 GATGACCGTACCGCCCAATGACCCGACAGCAAAATCGATGACTTCCGAGGAAGCT 11454
Qy 6239 GGTGCATATGCGACGCGCGCT-----TGAC 6265
Dh 11455 TGTGCATATATGATCAGCTGAGCTGTATATTAATGCCCGCTTACCGCGGCAATATAGCAAC 11514
Qy 6266 ACTAAACCTGATGATATTTCCGAGGAAGACAGTGCATTAATGCTGTGAGTGT 6318
Dh 11515 ACCAAACTCGAGTATTTCCGAGGAAGCGCAGTGCATTAATGCTGTGAGTGT 11567

RESULT 12
AAV60125
ID AAV60125 standard; DNA; 16656 BP.
XX
AC AAV60125;
XX
DT 25-MAR-2003 (updated)
DT 04-DEC-1998 (first entry)
DE
XX Representative eukaryotic layered vector initiation system.
XX
KM Eukaryotic layered vector initiation system; stimulate; immune response;
KM ss.
OS Synthetic.
OS Sindbis virus.
XX
PN US5814482-A.
XX
PD 29-SEP-1998.
XX
PF 30-OCT-1996; 96US-0739158.
XX
PR 14-SEP-1994; 94MO-US10469.
PR 15-MAR-1995; 95US-0404796.
PR 15-SEP-1993; 93US-0122791.
PR 18-FEB-1994; 94US-0198450.
PR 30-NOV-1994; 94US-0348472.
PR 18-JAN-1995; 95US-0376184.
XX
PA (DRIV/) DRIVER D A.
PA (DUBE/) DUBENSKY T W.
PA (JOLI/) JOLLY D J.
PA (POLO/) POLO J M.
PI Driver DA, Dubensky TW, Jolly DJ, Polo JM;
XX
DR WPI; 1998-541753/46.
XX

QY 2380 TTAAGACTTCAGAGCATTAAGAGGAGCCCAATCACCTCTAGCGGTGACCTAAATAG 2439
D 7543 AAAAGACATTCAGAGCATTCAGAGGAGAAATTAAGCATCTCTAGCGGTGCTAAATAGT 7602
QY 2440 TGACGTAGTACAGACGCACTAC-----CCACCGGCAAAATGTTTCCATACC 2486
D 7603 CAGCATGTACATTTCTACTAATTAATCTAACACACCAACCATGATAGAGGATTTCT 7662
QY 2487 CTCAGCTGAATTTTCCACCAATTTAACCTTAACCAATTCGATGGCTTACCGAATCCAAACC 2546
D 7663 TTAACATGTCTGGCCCGCCGCCCCCTTCGCGCCCCCAGCTGCATGTGAGAGCGGAGAA 7722
QY 2547 CTCCTAGAGCGCCGCTGAGAGCGGTTTCGCCCCCGCTGGCTGTCTCAATTCGAATCTTA 2606
D 7723 GGAGGACAGCGGCCCCCATGTCTCCCGCAACGGGCTGGCTTCTCAATTCAGACACTGA 7782
QY 2607 GGAGGTGATAGTCAACTTGAACCTTCAACAGATCACTAATCCGCGCAGGTGCA- 2665
D 7783 CCACAGCGGTGAGCCCTAGTCAATTGGACAGGCACTAGACCTCAACCCCGCTCAC 7842
QY 2666 -----CGCCAAAGAAAGAAAGAGTGTCTTAAGCCAAACCTAATCAGCTTAAAAAG 2720
D 7843 GCCCGCACCGCGCCAGAAAGAGAGGCGCCCAAGCAACCGAAGCGAAAGAAACCA 7902
QY 2721 AGAAGCAGCAAGCCAGAGAGAGCAAAAC---GCAAGCTTAAACAGGGAACGACACGTA 2777
D 7903 AAAAGCAGAGAAAGAAAGAAAGACACTTGCAAAACCAACCGGAAAGAGACAGCGCA 7962
QY 2778 TGTGTATGAGTGTGAGTGGACAAAGACATTTCCGATCATG--CTGAACGGCCAAAGTGA 2834
D 7963 TGGCACTTAAGTTGAGGCGCCACAGATTTGTGACGTCAAGAAAGAGAGCGAATGTCA 8022
QY 2835 ATGATATTCCTGGTGTGTGAGAGAGGCTGATGAACCACTCAAGTTAAAGAAA 2894
D 8023 TCGGGAACGCACTGGCCATGGAAGAAAGTAATGAACCTCTGCACTGAAGAAACCA 8082
QY 2895 TTGATTAATGAGCAATTAGCGCGCGTGAATTAAGAAAGGCTAGCATGTACGACTTGGAGT 2954
D 8083 TCGACACCCCTGTCTATCAAGGCTCAAAATTTACCAAGTGTCTGACATACGACATGGAGT 8142
QY 2955 AGCGGACGTTTCCCGAAGACATGAATTCAGACAGCTGCACTGACACGACCAAAACCAAC 3014
D 8143 TCGACAGTTGCGCATCAAGTGAAGTGAAGGATTCACCTAAGCCAGTAAACACCCCG 8202
QY 3015 CGGCTTCTAACAATGCGACACCGCGGAGTCCGATGAGAAATGGGAGATTTACCGTAC 3074
D 8203 AAGGATTTATTAATCTGGACCAAGAGCGGTGACGATATAGAGATTAATCAATCC 8262
QY 3075 CGAAGAGAGTGGCGGGAAGAGCGACAGCGGAACCGATCTGGAACAAAGAGGACAG 3134
D 8263 CTCGCGAGTATGAGAGGAGAGAGAGAGAGCGGTGCTGATGATTAATCCCGTGGG 8322
QY 3135 TTGTGGCTATTTCTAGAGAGTGCAAATGAGGCAACGCTGACGCGCTTTCAGTGTCA 3194
D 8323 TTGTGCGGATATGCTCGGTGGCGCTGATGAAGAAACGAATGCCCTTTCCGTCGCA 8382
QY 3195 CTTGGAACCAAGAGGGGTGACATTAAGGATTAACCCCGAAGTTCGAACCGGT--- 3251
D 8383 CTTGGAATAGTAAAGGAAAGCAATTAAGACGACCCCGAAGGAGCAAGAAAGTGTGTCG 8442
QY 3252 -----CACTAGTAAAGCGCTATGCGGTTCCTGAATGTACGTTCCCATGCAAAAC 3305
D 8443 CAGACACATGTGTACCGCAATGTGTGCTCGGAATGTAGCTTCCCATGCAACGCGC 8502
QY 3306 CACCGGTGTCTATTCATGACGCAAGACGAACACTGACGTCTCGAAGAGACGTGCG 3365
D 8503 CGCCCAATGTCTATACCGCAACCTTCAGAGCCCTGACATCTTGAAGAGACGTA 8562
QY 3366 ACATATCAAAATTAAGACAGCTGTGGAACGTTGAATGTCCA---TCAAGCGCGC 3422
D 8563 ACCATGAGGCTTACGATACCTGTCTCAATGTGCGGTGGAGTGTGTGCGCAAG 8622
QY 3423 CCAAGCAAGCATTAACGATGACTTCACTGACCAAGTCCCTACCTGGGGTTCTGCCGT 3482

D 8623 GAAAGAGAGCGTGTGAGACACTTACCTCTGACAGCCCTACTTGGGCAATGTCTGT 8682
QY 3483 ATTGCAACACTCAAGCCCGTGTTCAGCCCAATTAATAATTGAGAACGTGTGGAGCAAT 3542
D 8683 ACTGCCACCATATCGAACCGGTCTTCAAGCTTGTAAAGTGAAGAGGTCTGGGACGAG 8742
QY 3543 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3602
D 8743 CGAGACATTAACCATATGCGATACAGACTTCGCGCCAGTTTGGATAGACCAAGCGAG 8802
QY 3603 CTGCGATGTACCAAAATTCGTTTACATGTCTTTCGACCAAGCAACATGACATCAAGAG 3662
D 8803 CAGCAAGCGCAACCAAGTACCGCTACATGTCTTGAAGAGATCAACCGTTAAAGAG 8862
QY 3663 ACATATGAGAAATATGCTATGACACATCTGACACCTTCGCTGCTTGGCCACAAAG 3722
D 8863 GCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8922
QY 3723 GGTACTTCTGTAGCTCAATGTCTCTCAGATGACAGTGTAAACGCTGATGATGATGATGAT 3782
D 8923 GATCTTTCTCTCGCAAAATGCTCTCAGGAGACGCTTAAGCTATGATGATGATGATGAT 8982
QY 3783 GAGCATCTGAAATTCATGACCGGTGAGAAAGATCAGAGAGAGATTGTGCTGAGAG 3842
D 8983 GCAACTCAGCAACGTCATGATACATGCGCCGCAAGATTAACCAAAATTCGTGGAGCGG 9042
QY 3843 AGAGATCTTGTCTCCACCCGCTCCATGGAAGCTGTGTAAGTGCACGTTTACATCACT 3902
D 9043 AAAAATATGATCTACCTCCGCTTCAAGTAAAGAAATTCCTTGCAGGTACGACCGTC 9102
QY 3903 TGAAGAGAGCTGTGCGGGTATCAATCAATGACACAGGCGCCACAGCGCTATAGT 3962
D 9103 TGAAG---ACAATGACGCTCATCTATGACAGGCGCGGACCGCAAGCTTATCAT 9159
QY 3963 CCTATCTGAGAGAAAGCTGTAGCGGAAGTGTATTAACCACTTCTGCAAGAGAGCTCA 4022
D 9160 CCTACTGGAAGATCATCAGGGAAGTATACGAAGCCGCACTGTGGAAAGAACATTA 9219
QY 4023 CTTAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4082
D 9220 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9279
QY 4083 ACGGCTGCACTAAAGCAAAACAGTGTCTCTCAAGAGCGACCAACGAAATGGGCT 4142
D 9280 CTGCTGACCGCCATTAAGCAGTGTGCTCTATTAAGAGCGACCAACGAAATGGGCT 9339
QY 4143 TCAACTGCGCGATCTTATTAAGGACACAGACCACTCAAGTCAAGATTAATGCAATTC 4202
D 9340 TCAACTCAGCGACTGATGACATGACAGACCAACGCGCAAGGAAATTCGATTTGC 9399
QY 4203 CATTCGCTTGAACCGACAGTGTGCGCGGTTCCGTTAGCTCAACGCCATCAAGTCA 4262
D 9400 CTTTCAAGTTGATTCGCGGTGCTGATGATGATGATGATGATGATGATGATGATGAT 9459
QY 4263 AGTGTCTCAAGAGGATGACCTTCACTGATGATGATGATGATGATGATGATGATGATGAT 4322
D 9460 ATGGCTTAAACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9519
QY 4323 GAAATTTGGGCTGCGAGACAGACCAACAGCAGATGATTAAGGCTTATTCAGGA 4382
D 9520 GGAAGTATGAGGGAACCCCGAACCACCACTGAATGATGATGATGATGATGATGATGAT 9579
QY 4383 ATTTTCTGTGGGCGGAGAGAGGCTGATGATGATGATGATGATGATGATGATGATGAT 4442
D 9580 ACTTCAACGCTGACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9639
QY 4443 TGTGGGCGAGAGTGTGACACAGGCAACCAATGATGATGATGATGATGATGATGATGAT 4502
D 9640 TCTATGCGCAAGAGTGTGACACAGAGACCTTCAAGATGATGATGATGATGATGATGAT 9659
QY 4503 ACTATTAATCGGACCTCAAGTGTACATGATGATGATGATGATGATGATGATGATGAT 4562

Dh 9700 ATTACTACATCCGCTCTGTATACACCTTTAGCCGCTGCATCAGCTACCGTGGGA 9759
Qy 4563 TCTGTAGGACCTGATCATGACGAGCTTGATGCGCAAGCAAGAGACCTGCTGA 4622
Db 9760 TGATATTGGGCTAATGTTGACGTGTATATGCTTGTAAAGCGCGGTGCTCTGA 9819
Qy 4623 GCGCATACGCGCTTGCACCGAAACGGAACCCACAGCATTTACGCTTTTGTCTGA 4682
Db 9820 GCGCATACGCGCTTGCACCGAAACGCGTATCCAACTTCGCTGACCTTTGTCTGG 9879
Qy 4683 TTGCGCAACCAACGCTGAAACATTTGGAGAACTTTGACCATCTGTGTTAAACAC 4742
Db 9880 TTAGGTGCGCCAAATCTGAAAGTTCACGAGACCATGATTAATCTGTGTGAAAGTC 9939
Qy 4743 AACGCTTCTGTGGGACAGTTGTGATCTCTGCGACGCGTGTATATCTGCTCGGT 4802
Db 9940 AGCGCTTCTGTGGGACAGTTGTGATCTCTGCGACGCGTGTATATCTGCTCGGT 9999
Qy 4803 GCTTTTCAATGCTGATGCTTTTATTTATTTGTTGACAGCGCTGCTGGGGAAGTGAAG 4862
Db 10000 ACTGCTCTGCTGCTGCTGCTTTTATTTATTTATTTGCTGGGCGCTTACCTGGAAGTGAAG 10059
Qy 4863 CTTTGAACATGCGACCACTGTGCGCAATGTTCCGGGATCCGTTAAAGCGTGTGCG 4922
Db 10060 CTTAGGAACATGCGACCACTGTGCGCAATGTTCCGGGATCCGTTAAAGCGTGTGCG 10119
Qy 4923 AACGCGAGTTAGCGGACCTTAACCTGAGATACGCGTGTCTATCGGAATTAACAC 4982
Db 10120 AAAGGCGAGGATGATCCCGCTCAATTTGGAATACATGTCATGTCGAGGTTTTCG 10179
Qy 4983 CTTCAACTAACAGAGATACGATGACCTGCAATTCACACAGTATTCCTTACACACAAG 5042
Db 10180 CTTCAACCAACAGAGATACATTAATCTGCAATTAACACAGTATTCCTTACACACAAG 10239
Qy 5043 TTTAAATGCTGCGGCTGCTCGAGTCAAGAGCTCTCAAGGCGGATTAACAGTCCGCG 5102
Db 10240 TCAATGCTGCGGCTGCTCGAGTCAAGAGCTCTCAAGGCGGATTAACAGTCCGCG 10299
Qy 5103 TTTTGGGCGGTGTACCTTTATGTTGAGGAGGCGCAATGCTTCTGTGACATGGA 5162
Db 10300 TCTTGGAGGCGGTCAACCTTTATGTTGAGGAGGCGCAATGCTTCTGTGACATGGA 10359
Qy 5163 ACACACAATGAGTGAAGCGTACGTCGATGCTGTCACAGTCACTTAATGACAGGAG 5222
Db 10360 ACAGCAATGAGTGAAGCGTACGTCGATGCTGTCACAGTCACTTAATGACAGGAG 10419
Qy 5223 TCGCACTAAAGATTCAACAAGTCTCTGAAGTGGCGCTGCGTATGATTAACGGAACA 5282
Db 10420 AGGCAATTAAGGTGACACTGCGCGATGAAGTGAAGTGAAGTGAAGTGAAGTGA 10479
Qy 5283 CCACCGGCGACCTGATACGTTTGTCAATGGGTGACCGCAGTTCCTACGGGACCTGA 5342
Db 10480 CTACCAATTTCTTAATGATGTAGTGAAGGAGTGAACACGAGAACTGTTAAAGCTTGA 10539
Qy 5343 AGGTATAGCAGAGGCGGATATACAGCGCTTTTTCACCTTTGACCATTAAGTGTGATCA 5402
Db 10540 AAGTATAGTGAAGGCGGATATACAGCGCTTTTTCACCTTTGACCATTAAGTGTGATCA 10559
Qy 5403 GAAAGGCGCTTTTATCAACTACGATTCCTGATATGAGTGAAGTGAAGCAGAGCGT 5462
Db 10600 ATCGGCGCTGTTGATCAACTATGATCTCCGGAATATGAGCAGATGAAGCAGAGCGT 10659
Qy 5463 TCGGGAATTTCAAGATCTCTGCTGTGATGCTACAGACATTAAGTGAAGCAGAGCGT 5522
Db 10660 TCGGGAATTTCAAGATCTCTGCTGTGATGCTACAGACATTAAGTGAAGCAGAGCGT 10719
Qy 5523 GCGTGTAGAGCTTCTGTCAAGACATCAAGTCCCTTACACCGCAAGATTAAGGAT 5582
Db 10720 GCGTGTAGAGCTTCTGTCAAGACATCAAGTCCCTTACACCGCAAGATTAAGGAT 10779
Qy 5583 ATGAATGTGAAGAACATCTGAGACGACCTTGTCAAGAACACAGCACATTTGTGATGA 5642
Db 10780 TTGAGATGTGAAGAACATCTGAGACGACCTTGTCAAGAACACAGCACATTTGTGATGA 10839

Qy 5643 AAATTGAAGTGAAGCTTGTGAGGCTTAATGCTTAACTGCTTAAAGGACATCCCTATCTGA 5702
Db 10840 AGATTGCAATTAATCCCTTCCGAGCGGTGACCTGTTCTTATAGGGAACATTTCCATTTCTA 10899
Qy 5703 TTGACATCCCTGATGACACTTTTGTGAGATCATGAGATCAACAATTTTGAAGTTA 5762
Db 10900 TTGACATCCCTGATGACACTTTTGTGAGATCATGAGATCAACAATTTTGAAGTTA 10959
Qy 5763 GCTGACAGTGAAGACACTGCTATTTATTTGACAGCTTTGTGTTCTTAACTTACAGT 5822
Db 10960 AATGTGAAGTGAAGACACTTATTCAGACAGCTTTGCGGAGTGCACCTGCACT 11019
Qy 5823 ACAAGCTGACAGGAGGAGCATTTGTCAGTTCATCTCCACCTCCACGACGCTTTGA 5882
Db 11020 ATGATTCGACCGGAGAGTCAATGCCCCGTACATTCGATTCGACACAGCACTTCC 11079
Qy 5883 AGAAGCGACCAACATGTGACTGCGGTAGGACAGATTAACATTTTGAACATGCA 5942
Db 11080 AAGAGTGAACAGTACATGCTCTGAGAAAGAGCGGTGACAGTACATTTGACACCGGA 11139
Qy 5943 GCCCAAGCAAAATTTTATGTTGCTATGCGGACAGAGTCCACTGCAATGCTGAAT 6002
Db 11140 GTCCACAGCGCACTTATGCTATGCTGTGTGGAGAGAACACATGCAATGCAAGAT 11199
Qy 6003 GTAAACCAACGCGCGACACATTAATGAGAACACATTAAGTGAACCAAAATTCACAG 6062
Db 11200 GTAAACCAACGCTGACCATTAATGAGAACACATTAAGTGAACCAAAATTCACAG 11259
Qy 6063 CGGAGTTTCCAAACATCTTGAACCTGCTTGTGACCTTTTGGGAGACATCATCC 6122
Db 11260 CGGCAATCTCAAAACATCATGAGAGTGTGCTTTGCCCTTTTGGCGCGCTGTGCTG 11319
Qy 6123 TCAATTTTGAAGCTTATGATGTTGTGCTGACGCTTATCTTAATAACACAGTAT 6182
Db 11320 TATTAATTAATGAGCTTATGATGTTTGTGCTGACGATGATCTGACACAGAGAT 11379
Qy 6183 GACTGAGCGCGGACACTGA-CATAGCGTAAACTGATGATCTTCCGAGAGAGTGT 6241
Db 11380 GACGCTACGCCCAATGATGAGACAGCAAAACTGATGATCTTCCGAGAGAGTGT 11439
Qy 6242 GCATTAATGCAACGCGCGCT 6261
Db 11440 GCATTAATGCAACGCGCTGT 11459

RESULT 13
AAV42364
ID AAV42364 standard; DNA; 16656 BP.
XX
AC AAV42364;
XX
DT 02-OCT-1998 (first entry)
XX
DE Representative eukaryotic layered vector initiation system sequence.
XX
KM Representative eukaryotic layered vector initiation system;
XX DNA alphavirus; structural protein expression; inhibit; pathogen;
OS immune response; stimulate; ss.
XX
OS Sindbis virus.
XX
FN US5789245-A.
XX
PD 04-AUG-1998.
XX
PF 30-OCT-1996; 96US-0741881.
XX
PR 15-MAR-1995; 95US-0404796.
PR 15-SEP-1993; 93US-0122791.
PR 18-FEB-1994; 94US-0198450.
PR 30-NOV-1994; 94US-0348472.
PR 20-JAN-1995; 95US-0376184.

PR 30-OCT-1996; 96US-0741881.
XX (CHIR) CHIRON CORP.
XX Chang SMW, Driver DA, Dubensky TW, Ibanez CE, Jolly DJ,
PI Polo JM;
XX WPI, 1998-446089/38.
XX
PT DNA alpha: virus structural protein expression cassettes - for
PT producing recombinant alpha: virus particles
XX
PS Disclosure; Fig 3A-H; 140pp; English.
XX
CC The present sequence is a representative eukaryotic layered vector
CC initiation system derived from Sindbis virus. The specification describes
CC a DNA alphavirus structural protein expression cassette which comprises
CC an inducible promoter and an alphavirus structural protein gene, where
CC the promoter directs the expression of the alphavirus structural protein
CC gene upon induction of the promoter within a cell, and where prior to
CC induction within the cell, the expression cassette does not express
CC sufficient quantities of structural proteins to be cytotoxic to a BHK
CC cell containing the expression cassette. The products may be used to
CC inhibit pathogens and stimulate an immune response.
XX
SQ Sequence 16656 BP; 4467 A; 4349 C; 4183 G; 3657 T; 0 other;

Query Match 36.5%; Score 2369.2; DB 19; Length 16656;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3693; Conservative 0; Mismatches 1983; Indels 44; Gaps 9;

QY 580 ACTGACGGTATGAAGCGGAGCGGTATATTTCTCATCGGAAACAGGCCAACGCTTC 639
DB 5746 ACTGACCTAACCGGGGTAGGTGGTACATATTTTCAGCGACACAGGCCCTGGGCACTTGC 5805
QY 640 AACGAAATCAGTACGTCAATGTAACTACAGAACTATATTTGGATCGGCCCTGCATG 699
DB 5806 AAAAGAAAGTCGCTCTGAGAACGACGTTACAGAACGACCTTGGAGCGCAATGCTCGG 5865
QY 700 AGAAGTATTAAGCCCGCCGCTCATCTCGAAGAGAGAAATGTTACAGAAAGCTGC 759
DB 5866 AAAAATTCATGCCCCGCTGCTGACACGTGAAAGAGAACTCAAACTCAGGTACC 5925
QY 760 AATTATGGCCTCTGAAGGAATGAAGCAGTATCATACGAAAGTAGAATAATGA 819
DB 5526 AGATGATGCCACGAGCCCAAAAGTAGTACCACTCTGTAAGTAAATAATCAGA 5985
QY 820 AAGCAATTACAGCGAGCGACTATTTCTGATTGGGCAATATCTATCATCAGAAATGA 879
DB 5986 AAGCATATACCATGAGGAGTACTGTGAGACTACGACTGTATATCTGCGACAG-- 6042
QY 880 ATCTCTGAGTGTATACAGATCAATATCTGTACCAATCTACTCTGCAACGGTAATTA 939
DB 6043 ATCCAGCCAGAAATGCTATAGATCACTATTCGAAACCTATGTACTCCAGTGCCTACCG 6102
QY 940 ACAGGTTACATCTGCAGAGGCGGCTTAATAAGTCGCACTTATGTTATCCAGAGATT 999
DB 6103 CGAACTACTCCGATCCAGATTCGCTGTAGCTGTCTGTACCAATATTCATGATGAATCT 6162
QY 1000 ACCCTACAGTACGAGTATTTGTATTAACAGATGAATAGATCGTATCTTTCAGATGTGG 1059
DB 6163 ATCCGAGAGTAGATCTTATCAGATTAATGAGAGTAGATGCTTACTTGAATATGGTAG 6222
QY 1060 ACGGCGCATCTGCTGTGTATATACAGCACTTTTGTCCGGTAACTGAAAGCTAAC 1119
DB 6223 ACGGCAAGTGCCTGCTGTATATCTGCACCTTCTGCCCTTAAGCTTGAAGTTAAC 6282
QY 1120 CAAGAAGCATAGCATTTTGTAGAGAGATAGATGAGCGGTCCATCGGCTATACGA 1179
DB 6283 GGAAGAAACATGAGATAGAGCCCGGAATATCCGAGTGCCTTCAACAGAGTACGA 6342
QY 1180 ATACATTCAAAATGATTTGGCTGAGCTACTAAAGAAATTCGAACTTACCCAAATGC 1239
DB 1180 ATACATTCAAAATGATTTGGCTGAGCTACTAAAGAAATTCGAACTTACCCAAATGC 1239

DB 6343 ACACGCTACAAAATGTCATATTCGCCCAACTAAAGAAATTCGAACGTCACGAGATGC 6402
QY 1240 GAGAAATTAAGTCTTGAATTCGGCGGCAATTAATGATTTGTTTCAGAAATAGCAT 1299
DB 6403 GTAACTGCGCAACACGAGTACGACCACTTCAATGTGCAATGCTTTGCAAAATATGAT 6462
QY 1300 GCAATGATGATGACTGGGATACCTTTTCGCGTAAACCTATTCCGCTAACTACAGAAACG 1359
DB 6463 GTAAATGACGATATTTGGAGAGAGTTCGCTCGAAGCAATTAAGATTAACATGAGTTTG 6522
QY 1360 TTAACGAAATGTGACAAACCTGAAAGGCGGCAAGCAGCAGATTTGTTGCAATATCTC 1419
DB 6523 TCACCGCATATGTAGTAGTACGTAAGAGGCGCTTAAGCGCGCAACTATTTGCAAGAGCT 6582
QY 1420 ATATATCTAAACCGTTGACAGAGATATCCATGATGATCAATTCGTCATGATCTTAAGAG 1479
DB 6583 ATATATTTGTCCTCATTCGCAAGAGTGTCTATGATATGATTCATGATGATGATGAAAGAG 6642
QY 1480 ATGTCAAAATTAATCTCCGCGACGAAACATACAGAGAGCGGCTTAAGTGCAGATTATTC 1539
DB 6643 ACCTGAAAGTTACACGAGCAGCAACACACAGAAAGAACCGAAAGTACAAAGTATAC 6702
QY 1540 AGGCTGAGATCCCTTGTCTTACCGCTTACCTTTGCGGATTCATTCGGAATTAATCCGTA 1599
DB 6703 AAGCCGAGAAACCCCTGGCGACTGTACTTATGCGGAGTTACCGGGAATTAATGCTGTA 6762
QY 1600 GACTGAATGGGGTCTTTCGCAAAATATCCATCTCTTCGACATGTCGCGGAAGATT 1659
DB 6763 GGGCTTACGCGGCTTCTGCTTCAAACTTACACGCTTTTGAATGTCGCGGAGGAT 6822
QY 1660 TTGATGCAATTAATTTGCTGAACATTTCCACACCGCGCACCCAGTATTTGAAACGACATCG 1719
DB 6823 TTGATGCAATTAATTTGCTGAACATTTCCACACCGCGCACCCAGTATTTGAAACGACATCG 6882
QY 1720 CGTGCTTGAATAAAGCAAGCAGAGCTATCCCATTTGCGGCTGATATCTTGAAG 1779
DB 6883 CATCATTTGCAAAAGCAAGCAGAGCTATCCGTTAAACCGCTGATATCTTGAAG 6942
QY 1780 ACTTAGGTGCAACCAACGCTCTTATGATTTGATGAGAGGGGGTTCGGAATATCAT 1839
DB 6943 ACCGGGTGATGATCAACCACTACTGATTTGATGAGAGGGCTTTGGAAATATCAT 7002
QY 1840 CTGTGCACTACTACAGAAAGAGTTTAAATTTGGTGCATGATGAATCCGATATGT 1899
DB 7003 CCACCCATCTACTACGAGTACTCGTTTAAATTCGGGGGAGATGATGAAATCCGGAATGT 7062
QY 1900 TCTTAACGCTGTTTGTCAACACTATGCTCAATATCATGATTTGATGACAGATCTACG 1959
DB 7063 TCCCTCACTTTTGTCAACAGATTTTGAATGCTTATTCGCGCAGAGATCTGAAAG 7122
QY 1960 AAGGTTAAACAGTACGCGTGGCGGCTCTTATTCGCGCAGATTAATCATATGTCATGGT 2019
DB 7123 AGCGGCTTAAAGTCCAGATGTGACGCTTCAATGCGACGACACATATATCATGTAG 7182
QY 2020 TCGTCTCCGACACCTTGTAGCGGAGAGATGCCCACTTGGCTGAACATGAAAGTAAAA 2079
DB 7183 TAGATATGCAAAAGAAATGCTGAGAGGTGGCGCACCTGCTCAACATGAGATTAAGA 7242
QY 2080 TTAATGATGATTAATTTGATTAAGAACACCTTACTTCTGTGGGAGTTTATCTGTGTG 2139
DB 7243 TCATCGACGAGTATGCTGAGAGACCACTTACTTCTGCGGAGTTTATCTTGTGAG 7302
QY 2140 ACAGATTAACAGCAGCGCTGAGAGTGCACACCTCTAAAGAGCTTTTAAGCTTG 2199
DB 7303 ATTGCTTACTTCCACAGCGGTGCGGCTGAGATCCCTGAAAGGCTCTTTAAGTTGG 7262
QY 2200 GAAACCATTTGCGAGTATGATACCAAGATTCGACCGCGCGGACATGATGATG 2259
DB 7263 GTAAACCGCTCCAGCGGACGAGAGACAAACGAAAGACGAGCGCTCTGATGATG 7422
QY 2260 AAGCAATGCAATGAGACGAATTTGAATTAACGAGATTAAGTAAAGCCGTAGATCA 2319
DB 7423 AAACAAAGGCGTGGTTAGAGTAGGTAAACAGGCACTTAAAGAGTGGCGGTGACGACCC 7482

QY	2320	GATACGAGATCAATACTGCGAGGCTGATCATCAAGTCTCTGTCCAGCTTTAGCCGAAACGG	2379
Db	7483	GGTATGAGAGTGAACAAATATTTACACCTGTCTACTGGCAATTGAGAACTTTTGCCAGAGCA	7542
QY	2380	TTAAGACTTAAAGGCACTTAAGAGGGAGACCAATTCACCCTCTAGGCGTGCATTAATAGG	2439
Db	7543	AAAGGCACTTCCAGCCCACTCAGAGGGGAAATTAAGCATCTCTACGGTGGTCTTAAATAGT	7602
QY	2440	TGACGTAGTAGCAAGCACTAC-----CCACGGGAGAAATGTTTCCATACC	2486
Db	7603	CAGCATAGTACATTTCACTGTGACTTAATCTACAAACACACACCATGAAATAGAGATTTCT	7662
QY	2487	CTCAGCTGAACCTTTCACCAAGTTTACCTTCAAAATCCGATGGCTTACCGAGATCCAAACC	2546
Db	7663	TTAATCATCTCGGCGCGCCCTTCCGGGCCCCCACTGCATGTGGAGGCGCGGAGAA	7722
QY	2547	CTCCTAGGCGCGCTGGAGGCGCTTTGGGCCCCCGCTGGCGTGCCTAAATCGAATACTCTTA	2606
Db	7723	GGAGCGAGGCGGCCCCGATGCTCGCTGCCCCCAACGGGCTGGCTTCTCAAAATCCAGCACTGA	7782
QY	2607	GGAGGTCGATAGTCAACTTTGACCTTTCAAAACAAGTATCACTTAAATCCGCGCCAGAGTCCA-	2665
Db	7783	CCACAGCCGTCAAGTCCCTAGTCATGTGACAGGGAATAGACCTTAACCTTAACCCCAAGTCCAC	7842
QY	2666	-----CCGCCAAAGAAAGAAAGAGTGTCTTAAAGCCAAATCTTACGCTTAAAGAA	2720
Db	7843	GCCCCCAACCGGCCAGAAAGAGGAGGCGCCCAAGCAACACCGAAGCCGAAGAAACAA	7902
QY	2721	AGAAAGCAGCAAGCCAAAGAGGAGCGAAAC--GCAAGCTTAAACGAGGGGAAACGAACAAGTA	2777
Db	7903	AAAGCGAAGAAAGAAAGAAAGAACCACTTCGAAAAACCAAAACCGGAAGAGAACGCGCA	7962
QY	2778	TGTGTATGAAGTTGGAGTCGGAACAAGACATTTCCGATCATG---CTGAACCGCCAAATGA	2834
Db	7963	TGGCACTTAAGTTGGAGGCGCAGATGATTGTTGCAAGTCAAGAACGAGAGCGAAGATGTCA	8022
QY	2835	ATGATATATGCCCTGCTTTCGGAGGAAGCGCTGATGAACCACTCCACGTTGAAGGAAAA	2894
Db	8023	TCGGGCAAGCACTGGCCCATGGAAGGAAAGGTAATGAACCTCTCAGCTGAAGAGAACCA	8082
QY	2895	TTGATATATGAACAATTAAGCGCCCGGAAATTGAAGAAAGGCTAGAGATGACACTTTGGAGT	2954
Db	8083	TCGACCAACCTGTGTGTATCAAAAGCTCAAAATTTACCAAGTCTGTAGCATATGACATGAGAT	8142
QY	2955	ACGGGCAAGCTTCCCAAGAACATGAATCAGACACGCTGCACTAGACCAACGACAAACAC	3014
Db	8143	TCGCACAGTTCCAGTCAACATGAGAAATGAGGCAATTCACCTACCAACGATGAACACCCG	8202
QY	3015	CGGGTTTTCACATCTGGCAACCAAGCGCGCAATGTCAGATATGAGATAGGGAGATTTACGTTAC	3074
Db	8203	AAGGATTTTAAATCTGGGCAACCAAGGAGGAGTGAAGTATGAGAGTATTAATTAACATCC	8262
QY	3075	CGAGAGAGTGGGCGGGAAGGCGCAAGCGGAAAGCCGATCTCTGAGCAACAGAGGCGAG	3134
Db	8263	CTCGGGAGTATGAGAGGAGAGAGAACACGCGGTCTGTCCGATATGATGATTAATCTCGGTCGGG	8322
QY	3135	TTGTGCTATTGTTCTTAGAGGTGCAATGAGGCGACGCGTACGCGCTTTCACTGTGTCA	3194
Db	8323	TTGTGCGCATATGTCTCGTGGCGCTGTATGAAGAAACAGAACTGCTTTGCGTGTCTCA	8382
QY	3195	CTTGGAAACCAAAAGGGGTGACATTTAGGGATTAACCCCGAAGGTTCTGAACCGGTGT---	3251
Db	8383	CTTGGAAATGTAAGGGAAGAACAAATTAAGACGACCCCGGAAAGGGAACGAAGAGTGTGCG	8442
QY	3252	-----CACTAGTTACAGCGCTATGCGTGTCTTGAATGTACAGTCTCCCATGCGACAAC	3305
Db	8443	CAGACCACTGTGTACCGGCATATGTGTTGCTTGGAATATGAGGTTTCCATGCGACCGCC	8502
QY	3306	CACCCGTGTGCTATTCACTGACGCCAGAACGAACACTGACGTGCTCGAAGAGAACGTG	3365
Db	8503	CGCCCACTGCTAATACCGCGAACCTTTCACAGGCGCTTCGACATCTTTGAAGAGAACGTGA	8562

QY	3366	ACAAATCCAAATTTAGACACCGTCTGCGAGAACGTTGMAATGTCCA---TCAGCCGGCC	3422
Db	8563	ACCAATGAGGCGCTTAAGATACCTCTGCTCAATGCGCAATATTCGCGGTGGAGTCTGTGGCAGAA	8622
QY	3423	CCAAACGAAAGATTATCCGATGACTTCACTCGACAGTCCCTTACCTGAGGGGTCTGGCCGT	3482
Db	8623	GCAAAAGAGGCGTGTGACGACTTTTACCTTGACAGCCCTTACTTTGGGCACATGCTCGT	8682
QY	3483	ATTGCGACACTCAACGCGGTGTTTCAGCCCCAATPMAAAATTGAGAACGTGTGGACGAAT	3542
Db	8683	ACTGCAACCAATACGAAACCGTGCTTCAAGCCCTGTTAAGATGAGACAGGTCTGGGACGAAG	8742
QY	3543	CTGATGATGATTCGATTGAATCCAGGTCTTGCGACAAATTCGGCTCAATTCAGGACGCA	3602
Db	8743	CGGACGATPAAACCATATCGCATACAGACTTCGCGCCAGTTTGGTATACGACCCAAAGCCGAG	8802
QY	3603	CTGGGAGATGTCACCAATTCGCGTTCACATGTCCTTTGCGACAGACCAATGACATCAAGGAAG	3662
Db	8803	CAGCAAGCGCAAAACAAGTACCGCTTACATGTGCTTAAAGCAGGATCAACCGTTAAAGAG	8862
QY	3663	ACAGTATGAGAAATATAGCTTACGACATCTGGACCCCTGCGTCTTGCGCACAAG	3722
Db	8863	GCACCAATGATGACATCAAGATTAGACCTCAGAACCCGTGTGAAGAGCTTATGCTACAAAG	8922
QY	3723	GGAATCTCTGTAGTACATATGTCCTTCAGGTGACAGTGTAAACCGTCAGTATCAAGAGC	3782
Db	8923	GATACTTCTCTCTCGCAAAATGCCCCCTCAGGGGACACGTTAACGGTTAGCATAGTAGTA	8982
QY	3783	GAGCATCTGAAATTTCAATGCAACCGTGGAGAAAAAGATCAGAGAGGAATTTGTGCGTATAG	3842
Db	8983	GCAACTCAGCAACGTCATGTACACTGTGCGCGCAAGATPMAAACCAAAATTCGTGGGAGGG	9042
QY	3843	AGAGTACTTGTGTCCCAACCGCTCCATGGAAGAGCTGTAAAGTGCCAGTTTACATCACT	3902
Db	9043	AAAAATATGATCTACCTCCGTTACGCTTAAGATAAGAAATTCCTTGACAGTGTATGACCGTCT	9102
QY	3903	TGAAGGAGACGTCTGCCGGGTACATPACATGCAACAGGCGGACCCACAGCCGTATAGT	3962
Db	9103	TGAAA---ACAATCGACAGGCTACATCATGTATGCAACAGCGGGGACCGCAGCTTATATCAT	9159
QY	3963	CCTATCTGGAAGAAAGCGTCAGGCGCAATGTATCATTTAAACAACCTTCTTGCGAAGAACGTCA	4022
Db	9160	CCTACCTGGAAAGATCATCAGGAGAAATTTTAAACGAAAGCCGCATCTCGGAGAGAACATTTA	9219
QY	4023	CCTACGATGTAAATGTGCGGCGATCAACAGCAACAGGTATCGTAGAGCAGCGCAACGAATGA	4082
Db	9220	CGTATGAGTGCMAATGTGGCGCATTAACAAGCCGAAACCGTTTGAACCCGACCGGAATATCA	9279
QY	4083	ACGCGCTGACTPMAAGCAAAACAGTGCATTGCTTACAGAGACGACCAAAACGAAATGGGCT	4142
Db	9280	CTGGTTGACCGGCATCAAGCAATGCGGTCTGCGCTTAAAGACGACCAACGAATGGGCT	9339
QY	4143	TCAACTGCGCGGATCTTATTTAGGCAACAGACCACTCAGTGCAGAGTTAAATTGCAATTC	4202
Db	9340	TCAACTCACCGCACTTATCATGACATGACACACACAGCGCCCAAGGGAAATTTGATTTGC	9399
QY	4203	CATTCCGCTTGACACGACAGTCTGCCCGGTTCCGTTAGTCTCACGCCCTTACAGTCCGA	4262
Db	9400	CTTTTCAGATTGATCCCGGGTGCCTGCAATGGTCCCTGTGGCCACAGCCCGCAATGTAATAC	9459
QY	4263	AGTGATTCAAAGGATCACCTTCCACTGTGACTGCAATGACACAAACATTTGCTGCAACGA	4322
Db	9460	ATGGCTTTAAACATCATGAGCTTCCAAATTAATATACAGACCACTTGAATTCCTACCCCA	9519
QY	4323	GAAATTTGGGGCTGCGAGCAGACGCAACAGCAGATGATTTACAGGTTCTATCTCAGGA	4382
Db	9520	GGAACATAAGGGGCAAAACCCGGAACCAACCACTGAATGATCTGTGGAAGAACGCTGAGAA	9579
QY	4383	ATTTTCTGTGGGGCGAAGAGGCTGGAGATACGATTTGGGGTTAACATGAAACGATGAGAG	4442
Db	9580	ACTTCACCGTGCACCGAGATGAGCTTGGAAATACATATGGGAAATCATGAGCCAGTAGAGG	9639
QY	4443	TCTGGGGCCAGAGATCCGCAACAGGCGCAACCCACATGATGTGGCGCATGAGATCATATCC	4502

D	b	9640	TCATATGCCCAAGATCAGACACAGAGAACCTTCACGGATAGGCCACACGAAATAGTACAGC	9699
Q	y	4503	ACTATTATCAATCCGCAATCCAGTCTACAGTGTCAATTGTGCTGTGTGTGTGCTCTTGTGTA	4562
D	b	9700	ATTACTACCAATCCGCATCTGTGTGACACCAATCTTAGCCGTGCATCAGCTACCGTGGCGA	9759
Q	y	4563	TCCGTGTAGGCACTGCATCAATCAAGCAAGCTTGCAATCCGCCAAAGAAAGAAAGACTGTGCTTA	4622
D	b	9760	TGATGATTTGGCGTAACTGTTGACAGTGTTAATGTGCTGTAAAGGCGCCCGTAAGTGTGCTTA	9819
Q	y	4623	CGCCATACGGGCTTGCACCGAAACGCAACGGTATCCACAGACTATAGCGGTTTTGTGTGCA	4682
D	b	9820	CGCCATACGCCCTGTGCCCCCAAGCCGTAATCCCACTTGCTGGGCACTTGTGTGCTGG	9879
Q	y	4683	TTGCGCCAAACCAACGCTGAAACAATTGGAGAAACTTTGAACCACTGTGTGTTTAACAACC	4742
D	b	9880	TTAGTGTGGCCCAATGCTGAAACGTTACCGAGACATGATGATTACTGTGTGCAACATG	9939
Q	y	4743	AACGCTTCTCTGTGGGCAAGTGTGCAATTCCTCTGGCAGGCGCTGTATATTCTGTCCGCT	4802
D	b	9940	AGCGGTCTCTGTGGGTCCAGTGTGCAATCTTTTGGCCGGGTTCATGCTTCAATATCCGT	9999
Q	y	4803	GCTTTATGCTGATGATGCTTTTATTTATTTGGTGAAGGCGCTGTGCTGGGGAAAGTGAAGC	4862
D	b	10000	ACTGTCTCTGTGCTGCTCTCTTTTATGTGTGTGCGGCGCTTACCTGGCGAAGTGAAGC	100585
Q	y	4863	CCTTCGAACATGCGAACCACTGTGCAAAATTTCCGGGGAATCCCGTATTAAGCGTTGGTGC	4922
D	b	10060	CTTACGAACATGCGAACCACTGTTCCAAAATGTGCAAGATACGTAATTAAGGCACTTGTGG	10119
Q	y	4923	AAGCGCAAGTTACGCGCCCACTTAACCTGGAGATCACGGTGTCTCATGGGAATTAAAC	4982
D	b	10120	AAAAGGCAAGGATTAAGCCCCCTCAATTGGAGATCACTGTCAATGCTCTGTGAGAGTTTTC	101797
Q	y	4983	CTTCAACTTAACAAGGATAGGTGACCTGGCAAAATTCACAACAGTATTCCTTACCAACAAG	5042
D	b	10180	CTTTCACCAACCAAGATTAATTAACCTTGCAAAATTCACACATGTGTGTCCCTCCCAAAA	102389
Q	y	5043	TTAAATGCTGCGGATCCCTCGAGTGCAAAGGCATCTTCAAAAGCGGATTAACATGTCCGCG	5102
D	b	10240	TCAATGCTGCGGCTCTTGGGAATGTAGCGCGGCGCTCATGCAACTATACCTGCAAG	102998
Q	y	5103	TTTTTGGCGGTGTGTACCTTTTCAATGTGGGAGGCGCAATGCTTGTGTGACATGTGA	5162
D	b	10300	TCTTCGAGAGGAGTCTACCCCTTATATGTGGGAGAGGCGCAATGTTTTTGGACATGTGA	103588
Q	y	5163	ACACACAACGTAGTGAAGCGGTAGCTGCAATTCGCTCAGACTGTCACTATATAGATCAGGAG	5222
D	b	10360	ACACCCAGATGAGTGAAGGCGGTACGTCAAAATGTTCAGACAGATTCGCGTCTGACACAGCGC	10419
Q	y	5223	TCGCATCTAAAAGTTTCAACACAGTGTCTGAAAGTCCGCTGCCTATATAGTACGCAACA	5282
D	b	10420	AGCGGATTTAAGTGTCAACTGTCCGCAATGAAGATAGGACTGTGCTATATGTGTACGGGAAC	104797
Q	y	5283	CCACCGCGCACTTGATAGCTTTGTTCAAATGGCGTCAACGCGAGTTCTCAACGGGACTGA	5342
D	b	10480	CTACCAATTTCTATGATGTGTACGTGAACGAGTACACACAGGAAGCTCTAAAGACTTGA	105398
Q	y	5343	AGGTCTATAGCAGGCGCCATATCAGCCGCTTTTACCCCTTTGACATATAGTGTGTATCA	5402
D	b	10540	AAGCTATAGCTGAGCAAAATTCACAGCATGTTTACGCCATTCATATAGTGTGTATTC	105998
Q	y	5403	GAAGGGGCTTGTTTACAACATCAAGCTTCCCTGATGTGAGGTATGAAGAACGAGAGCGT	5462
D	b	10600	ATCGCGGCTGTGTATCAACTATATGACTTCCCGGATATGAGCGGATGAAGCAAGAGCGT	106588
Q	y	5463	TCGCGATATTTCAAGCACTCTCGCTTGATGCTACAGACATAGTAGCCCGCATGACATAC	5522
D	b	10660	TCGGAGCATTTCAAGCTATACCTCTTGTGACTGAACAGATTCATATCGGACAGACGATTA	10719
Q	y	5523	GGCTGTGAAGCTTTCTGTCAAGACATCAAGTCCCTTACACCCCAAGCATATCAAGGT	5582

Db	10720	GGGTACTCAAGCCTTCGCCAAGAAAGTGATGATGTCCTGGTACAGACGACGAGCGCATCAGAT	107719
Qy	5583	ATGAAATGTGGAAGAACCACTACGAGACGACCCCTGCAAGAAACAGACCACTTTGGATGTA	5642
Db	10780	TTGGAGATGTGGAAGAAACCACTCAGGCGGCCCACTGAGAGAAACCGACCTTTCGGGTGTA	108339
Qy	5643	AAATGTGAAGTGAGCCTCTGCGAGCGCTTAACTGTCTTAACGGGACATCCCTATCTCGA	5702
Db	10840	AGATTGCAATGTAATTCGGCTCCGAGCGGTGAGACTGTTCATACGGGAACATTTCCATTTCTTA	108939
Qy	5703	TTGACATCCCTGATGACGCTTTTGTGAGATCATCAGAAATCACCMAATTTTGAAGTTA	5762
Db	10900	TTGACATCCCGAAGCGTGCTTATCAGGACATCAGATGACACACTGTCTCAACAGTCA	109559
Qy	5763	GCTGCACAGTACGACAGACTGCTATTATTTCTGCAGACTTTGTGTGTTCTTAACATTACAT	5822
Db	10960	AAATGTGAAGTCAAGTGATGTGCACTTATTCAGCAGACTTCGCGGGAGTGGCACCTGCACT	110119
Qy	5823	ACAAAGGTGACAGGAGGGGCAATTGTCACAGTTCACTCCCACTCCAGACAGCGTGTTTGA	5882
Db	11020	ATGTATCCGAGCGCGAAGGTCAAATGCCCTGTACATTGCGATTTCGACACAGCAACTCTCC	110799
Qy	5883	AGGAGCGACACACATGTGACTGCCGTAGGCGAGCATAACTACATATTTTACACATCGA	5942
Db	11080	AAAGTGTGACAGTACATGTCCTGAGAAAGAGCGGTGACAGTACACTTTAGACCGCGA	111399
Qy	5943	GCCCAACAAGAAATTTTATATGTTTCCCTATGCGGCAAGAAATTCACCTGCAATGCTGAAT	6002
Db	11140	GTCACACAGCGGAACCTTATGTTGCTGTGCTGTGTGGGAAGAAAGACATCATGCAAGAAAT	111999
Qy	6003	GTAACACACCGCGCGACCATATATTGGAGAAACCATATAAGTGCAGCAAGAAATTCAGG	6062
Db	11200	GTAACACACAGCTGACCATATCTGAGACACCCGCAAAAATGACCAAGAAATTTCAAG	112559
Qy	6063	CGGACGTTTCCAAAACATCTTGGACATGCGCTGCTTGCACCTGTGTTGGGGAGGATCATCC	6122
Db	11260	CCGCACTCAAAAACATCATGGAAGTTGGCTGTTTTCCTTTTCGCGCGCGCTGCTGCC	113119
Qy	6123	TCAATGTTGTAGAGCTTATATGTTGTTGTTGCTGACGCTCTATGCTTATAACACAGTGAAT	6182
Db	11320	TATTAATTATAGAGCTTATATGATTTTGTCTTGACAGCATGATGCTGATACACAGAAAGAT	113799
Qy	6183	GACTGAGCGGGGACACTGA-CATAGCGGTAAACGTGATGTACTTCGAGGAAGGTGGT	6241
Db	11380	GACGCTACGCCCCCAATGATCCGACGAAACATGATGTACTTCGAGGAACGTGATGT	114399
Qy	6242	GCATTAATGCCAGCGCGCGCT 6261	
Db	11440	GCATTAATGCAATCAGGCTGGT 11459	
RESULT 14			
AAV70684			
ID	AAV70684	standard; DNA; 16656 BP.	
XX	AAV70684;		
AC			
XX	20-MAR-2003	(updated)	
XX	DT	02-FEB-1999	(first entry)
XX	DE	Representative eukaryotic layered vector initiation system.	
XX	KW	Eukaryotic layered vector initiation system;	
XX	KW	alphavirus vector construct; gene therapy; ss.	
OS	Synthetic.		
OS	Sindbis virus.		
XX	US5843723-A.		
XX	US5843723-A.		
XX	01-DEC-1998.		
XX	30-OCT-1996;	96US-0739167.	
XX	PF		

XX 15-MAR-1995; 95US-0404796.
PR 15-SEP-1993; 93US-0122791.
PR 18-FEB-1994; 94US-0198450.
PR 30-NOV-1994; 94US-0348472.
PR 20-JAN-1995; 95US-0376184.
XX
PA (CHIR) CHIRON CORP.
XX
PI Belli BA, Chang SMW, Driver DA, Dubensky TW, Ibanez CE;
PI Jolly DJ, Polo JM;
XX WPI; 1999-044581/04.
XX
PT Alphavirus vectors constructs containing a 5' promoter of viral cDNA
PT by in vitro transcription - used in gene therapy
PS
XX Disclosure; Fig 3A-H; 140pp; English.
XX
CC The present sequence represents a representative eukaryotic layered
CC vector initiation system derived from Sindbis. The sequence exemplifies
CC the invention. The specification describes alphavirus vector constructs.
CC These constructs comprise a promoter 5' of viral cDNA which initiates
CC the synthesis of RNA from the viral cDNA by in vitro transcription,
CC followed by a 5' sequence which initiates transcription of alphavirus
CC RNA, followed by a nucleotide sequence encoding alphavirus nonstructural
CC proteins, a viral junction region which has been inactivated such that
CC viral transcription of a subgenomic fragment is prevented, an internal
CC ribosome entry site or a sequence which promotes ribosome read through
CC between adjacent reading frames, and an alphavirus RNA polymerase
CC recognition sequence. The recombinant alphavirus vectors can be used
CC for gene therapy.
CC (Updated on 20-MAR-2003 to correct PR field.)
XX
SQ Sequence 16656 BP; 4467 A; 4349 C; 4183 G; 3657 T; 0 other;
Query Match 36.5%; Score 2369.2; DB 20; Length 16656;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3693; Conservative 0; Mismatches 1983; Indels 44; Gaps 9;
QY 580 ACTGACGTTATGAAACGGGACGTTATTTCTCATCGGAAACAGGCCAAGCTCCTTC 639
DB 5746 ACTACTTAAACCGGGGTAGTGGGTACATATTTTTCAGCGACAGAGCCCTGGGACCTTGC 5805
QY 640 AACGAATCAGTACGTACATATGTAACATCAAGAACCTATATGATGGGCCCTCCATC 699
DB 5806 AAAAGAAAGTCCGTTCTGAGAACCAAGCTTACAGAACCCAGCTTGAGGCGCAATGCTCTGG 5865
QY 700 AGAAGTATTACGCCCCCGGCTCGATCTCGAAAGAGAAATGTTACAGAAAGAACTGC 759
DB 5866 AAAGATTCATGCCCCGCTGCTGACACGTGGAAGAGAACTCAAACTCAGGTACC 5925
QY 760 AATTATGCGCTCTGAAAGAAATAGAGACAGGTATCAATCAAGAAATAGAAAATATGA 819
DB 5926 AGATGATGCCCAAGAACCAAAAGATAGTACAGTCTGTAAGATAGAAATATCAGA 5985
QY 820 AAGCAATTACGCGGAGGAGTCACTTTTGTGATTTGGGACATATATCATCAGAAATGA 879
DB 5886 AAGGCATTAACCACTGAGGAGTACTGTGAGACTAGCACTGTATTAATCTCCACAG--- 6042
QY 880 ATTCCTGAGTGTATACAGATCAATTAATCTGTACATCTCACTGCAACGGTAATTA 939
DB 6403 ATCAGCCGAGATGCTATTAAGATCACTTATCCGAAACCTTTGTACTTCGATGCGTACCGG 6102
QY 940 ACAGGTTTACATCTGACAGAGTGGCGGTTAAACGTCACTTAAGTTATCCAGAGAAAT 999
DB 6103 CGAACTACTCCGATCCACAGTTCGCTGAGCTGTGTATCAACATATCTGCATGAGAACT 6162
QY 1000 ACCCTACAGTACAGTATATGTATTAACAGATGAATACAGATGCGATCTTACATGCTGG 1059
DB 6163 ATCCGACAGTACATCTTATACAGATTAAGTACAGATGAGTCTTAATCTGATATGAGTGG 6222
QY 1060 ACGGCGCATCGTGTCTGTATGATACAGCCACTTTTTCGCGCTAACTGAGAAAGCTACC 1119

DB 6223 ACGGACAGTGCCTGCTGATATCTGCACTTCTGCCCCGCTAAGCTTGAAGTTACC 6282
QY 1120 CAAAGAACATAGCTATTTTTCAGCCAGGATTAAGATCAGCCGCTCCCATGCTTATACGA 1179
DB 6283 CGAAAGAACATAGATATAGACCCCGGATATCCGAGTGGCTTCCATCAGCAGTACGA 6342
QY 1180 ATACATTACAAATATGATTTGCTGAGCTATCAAAAGAAATGCAAGTTACCCAAATGC 1239
DB 6343 ACAGCTACAAATATGCTCTATCTGCGCACTTAAAGAAATTTGACATCTCAGCGATGC 6402
QY 1240 GAGATTAACCTGTCTTGAATTCGCGCATTTAAATGTTATGTTTCAAGAAATCCGAT 1299
DB 6403 GTGAACCTGCAACATGGAAGTACAGGACATTCATATGTAATGCTTTGAAAAATATGCAT 6462
QY 1300 GCAATGATGATCTGGGATACCTTTGCGGATTAACCTTATTCGCTTACATACAGAAACG 1359
DB 6463 GTAATGACGATATTTGGAGAGATTCGCTCGGAAGCCAAATTAAGATTAACCACTGATTTG 6522
QY 1360 TTACGCAATATGTACAAAGCTGAAGGGCCGAAAGCAGCAATGTTTGGAAATATCTC 1419
DB 6523 TCACCGCATATGATGCTAGACTGAAGGCTTAAGCCGCACTATTTTGGAAAGAGCT 6582
QY 1420 ATAAATCAAACCGTTGCGAGAGATACCAATGATCAATTCGTCATGATCTAAAGAG 1479
DB 6583 ATAAATTTGCTCCATTTGCAAGAAAGTGCCTATGATAGTTGTGATGACATGAAGAAAG 6642
QY 1480 ATGTCAAAATTAATCTCCGCGCAGAAACATACAGAGAGCGGCTTAAGGTGAGATTTC 1539
DB 6643 ACGTGAAGATTAACACAGGCGCAGAAACACAGAAAGAAAGCCGAATTAAGTGAATAC 6702
QY 1540 AGGCTGACATCCCTGCTACACGCTTACCTTCCGGAATCCATCGGGAATTAAGCTGTA 1599
DB 6703 AAGCCGCAACCCCTGCGCACTGCTTATATGAGGATTAACCGGGAATTAAGTGCCTA 6762
QY 1600 GACTGAATGCGGCTCTTTCGCAATATTCATCTCTTTCGATGTCACGCGAAGATT 1659
DB 6763 GCGTTAAGGCGGCTGCTGCTTCCAAACATTCACAGCTTTTGAATGTCGCGGAGATT 6822
QY 1660 TTGATGCGATTAATTCGTAACATTTTCCACACGCGGACCCAGTATTTGAAACGACATCG 1719
DB 6823 TTGATGCAATATGACAGAACATCTCAAGGCGACCCCGTCTGAGAGCGATATCG 6882
QY 1720 CGTGGTTGATTAAGGGAAGACAGCGTATCGCATTTGGCGTGTATGATTCCTTGAG 1779
DB 6883 CATATTCGACAAAAGCCAGAACGATATGAGGCTTAACCGGTCTGATGATCTTGAGG 6942
QY 1780 ACTTAGGTGACCAACCGCTCTTAAATTTGATAGAGGCGGCTTCGCAATATCAAT 1839
DB 6943 ACTGGGTGTGATCAACCACTACTGACATTTGATGAGTGGCCCTTTGAGAAATATCAT 7002
QY 1840 CTGTGCACTTACGTAAGAAAGAGTTTAAATTTGCTGCATGATGAATCCGATATGT 1899
DB 7003 CCACCCATCTTACCGGGTACTCGTTTAAATTCGGGGCGATGATGAATCCGGAATGT 7062
QY 1900 TCTTAAGCGCTTTTGTAAACACTAGTCAATATCATGATTTGTACAGAGATTAACACTG 1959
DB 7063 TCTTCACTTTTGTCAACACAGTTTGAATGTCTTATGCAAGAGATTAACACTGAG 7122
QY 1960 AACGTTAAACCAAGCTGACGCGCGGCTTATTCGCGGAGATTAACATATGATGATG 2019
DB 7123 AGCGGCTTAAACGTCAGATGTGACGCTTATTTGGGAGACAGAACATCATATGAG 7182
QY 2020 TCGTCTCCGACCTTGAATGCGGAGAGATGCGCACTTGGCTGAACATGGAATTAATA 2079
DB 7183 TAGTATCTGAAGAAATAGCTGAGAGTGCACCACTGCGCTCAACATGAGAGTTAAGA 7242
QY 2080 TTATTTGATGAGTATATGATTAACAGCACTTCTCTGTGGGAGATTTATCTGCTGG 2139
DB 7243 TCATCGAGCGAGTATGCTGAGAGACCACTTACTTCTGCGGAGATTTATCTTGCAAG 7302
QY 2140 ACCAGATTAACAGGACAGCGCTGACAGTTCGAGACCTCTTAAAGGCTTTTAAAGCTTG 2199

Dh 7303 ATTGGTACTTCACAGCGTGGCGGCGGATCCCTGAAAAAGGCTGTTTAAGTTGG 7362
Qy 2200 GAAACCATTTGGCAGTGGATGATACCCAAAGCTGGACCCCGCGGGGCACTGATATG 2259
Db 7363 GTAAACCGCTCCAGCGAGCAGACGAAAGCAGAAAGCGGCTCTGCTAATG 7422
Qy 2260 AAGCAATGCGATGGAAAGATTTGAATTTACGACGATTTAGTAAAGGCTGATATCCA 2319
Db 7423 AAACAAAGGCGTGGTTTAAAGTAAAGTAAAGGCACTTTAGAGTGGCGGTGACGACC 7482
Qy 2320 GATACGATCATATCTGGCAGGCTGATCATCAGTCTCTGTCCACCTTAAAGCG 2379
Db 7483 GGATAGGTGAGCAATATTTACCTGCTACTGCGATTGAGAACTTTGGCCAGAGCA 7542
Qy 2380 TTAAAGACTTCAAGAGATAAGAGGAGGCCCAATCACCCTCTACGCTGACCTTAAATAG 2439
Db 7543 AAAGAGATTTCAGACCATAGAGGGGAAATTAAGCATCTCTAGGAGTCTTAAATAGT 7602
Qy 2440 TGACGTAGTACACGCACTTAC-----CCACCGGAGAAATGTTTCATACC 2486
Db 7603 CAGCATAGTACATTTCACTGACTAATATCAACACACACCATGAAATAGAGATTCT 7662
Qy 2487 CTGACGTAACTTTCCACAGTTTACCCTTCAAAATCCGATGGCTTACCGAGATCAAAAC 2546
Db 7663 TTAAATGCTCGGCGCGGCGGCTTCCGGGCGCCACTGCGCATGTGAGGCGCGAGAA 7722
Qy 2547 CTCTAGAGCGCGCTGGAGCGGCTTGGCGCGCGCGCTGCTGCTCAATGGAATTTA 2606
Db 7723 GAGAGCGGCGGCGCGCGCTGCTGCGCGCAACGCGGCTGCTTCAAAATCCAGCACTGA 7782
Qy 2607 GAGAGTGCATATGCTTCACTTTCAAAACAGTACCTTAATCCGCGCGCAGTCCA- 2665
Db 7783 CCACAGCGCTCAGTGGCTTGTATTTGGAACAGGCACTAAGACCTCAACCCCGAGTCCAC 7842
Qy 2666 -----CCGCCAAAGAGAGAGAGTCTCTTAAAGCAAACTTAAGCTTAAAGAA 2720
Db 7843 GCCCGCACCGCGCCAGAGAGAGCGGCGCCAGCAACCAACCGGAGAAACCA 7902
Qy 2721 AGAAGCGAGAGCCAAAGAGAGCGAAAC---GCAAGCTTAAACGAGGAAACGAAAGTA 2777
Db 7903 AAACGCGAGAGAGAGAGAGAGACCTGCAAAACCGGAAAGAGAGAGCGCA 7962
Qy 2778 TGTGTATGAAGTTGGAGTTCGACAAACATTTCCGATCATG---CTGAACGCGCAATGA 2834
Db 7963 TGGCACTTAAGTTGGAGCGCGACAGATTTGTCGACGTCAGAAAGAGAGAGAGATGCA 8022
Qy 2835 ATGAGATATGCTGCTGCTGGTGGAGAGAGCTGATGAAACCACTCCAGCTTGAAGAAAA 2894
Db 8023 TCGGGCAGCAGACTGGCCATGAGAGAAAGTAATGAAACCTCTGCACTGTAAGAAACCA 8082
Qy 2895 TTGATATGAAGCAATTAGCGGCGGTGAATTTGAAGAAAGGCTAGACATGTAGACTTGGAGT 2954
Db 8083 TCGACCACTGCTGCTATCAAAAGCTCAAAATTTACCAAGTGTGACGACATGAGATGAGT 8142
Qy 2955 ACCGCGACGTTCCCGCAGAAACATGAATCAGACGCTGAGTACAGACGAGCAAAACAC 3014
Db 8143 TCGGACAGTTGGCAGTACATGAGAGTGAAGGATTCCTTACACAGTGAACACCCCG 8202
Qy 3015 CGGGCTTCTACACTGCGACCAAGCGCGAGCTCAGATATGAGATGAGATTTAACGTTAC 3074
Db 8203 AAGGATTTCTAATCTGGCACCAACGAGCGGTGAGTATGAGAGTATTTACCATCC 8262
Qy 3075 CGAGAGAGTGGGCGGAG 3134
Db 8263 CTCGCGAGTGAAG 8322
Qy 3135 TTGTCGCTATTTGTTCTAGAGAGTCAAAATGAGGCGACGCGTACGCGGCTTTCACTGTTCA 3194
Db 8323 TTGTCGCGATAGTCTCGGTGGCGCTGATGAAGAGACGAACTGCGCTTTCCGTCGTCAC 8382
Qy 3195 CTGGAACCAAGAAAGGGGTGACATTAAGGATACCCCGAAGGTTCTGAACCGTGT-- 3251
Db 8383 CTTGGAATAGTAAAGGAGAGCAATTTAAGACGACCCCGGAGAGGAGACAGAAAGATGTCG 8442

Qy 3252 -----CACTAGTTACAGCGCTATAGCGTGTTCGATGTGACAGTTCGCCAGCAAAAC 3305
Db 8443 CAGACCACTGTGTACAGGCGCATGTGTTGCTCGGAAATGTAGCTTCCATGCGACCGCC 8502
Qy 3306 CACCGGTGTCTATTTCACTGACCGCAGAAAGCACTGCACTGCTGGAAGAGAGCTGCG 3365
Db 8503 CGCCCATGTCTATACCCCGGAACCTTCCAGAGCCCTCGACATCTTGAAGAGAAAGTGA 8562
Qy 3366 ACAATCCAAATTTACGACGCTGCTGGAAGAGCTTTGAAATGCCA---TCAAGCGCGC 3422
Db 8563 ACCATAGGCGTACGATCCCTGCTCAATGCCAATTTGCGGTGCGGATGCTTGGCAGAA 8622
Qy 3423 CCAAGAGCATTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3482
Db 8623 GCAAAAGAGAGCGTGTGAGACATTTTACCTGACAGCCCTTACTTGGGACATGCTCTGT 8682
Qy 3483 ATTGACACACTGACGCGCTGTTTACGCCCAATTAATTTGAGAAAGTGTGGAAGAT 3542
Db 8683 ACTGCCACATCTGAAACCGTGTCTGAGCGCTGTGATGATGAGAGGTGTGGGAGCGAG 8742
Qy 3543 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3602
Db 8743 CGAGCATTAACACATACGATACAGACTTCCGCCAGTTTGGATACGACAAAGCGGAG 8802
Qy 3603 CTGCGATGTCACCAAAATTCGTTACATGCTTTGACACGACGACATGATCAAGAGAG 3662
Db 8803 CAGCAAGCGCAACAGTACCGCTTACATGTCGTTAAGAGAGATCAACCGTTAAAGAG 8862
Qy 3663 ACAATGAGAAATATGATATGAGACATCTGACATCTGACCTGCTGTTGSCCAAAAG 3722
Db 8863 GCACATGATGATCATCAAAATTTAGCACTTCAAGAGCGGTGTAGAGCTTAAAG 8922
Qy 3723 GGTATCTCTGTTAGTCAATGTCCTCCAGGTGACAGTGTAAACCTGATATCAAGAGCG 3782
Db 8923 GATATCTTCTCCGCAAAATGCTTCAGAGGAGACGCTGAAGGTTAGATGTGTGTA 8982
Qy 3783 GAGCATCTGAGATTAATGATGACCGTGAAGAAAGATCAGAGAGAAATTTGTCGTAGAG 3842
Db 8983 GCACATGAGCAAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9042
Qy 3843 AGAGATATCTTGTCCACCGGCTCATGAGAAAGCTGTGAAGTCCAGCTTACGATCACT 3902
Db 9043 AAAAATATGATCTACCTCCGTTCAAGGTAAGAAATTCCTTGCACAGTGTACGACGCTC 9102
Qy 3903 TGAAGAGAGTGTGCGGCTGATATCAATGATGACAGCGCCAGCCACAGCTTATAGT 3962
Db 9103 TGAAG--ACAACTGACAGGCTACATCATATGACAGGCGGAGACGCAAGCTTATACAT 9159
Qy 3963 CTTATCTGAGAGAGCGTCAAGGAGAGTATTAATTAACCACTTCTGCAAGAGAGCTCA 4022
Db 9160 CTTATCTGAGAGAGATCATGAGGAAATTTACGCAAGCGCCATCTGAGAAAGACATTA 9219
Qy 4023 CTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4082
Db 9220 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9279
Qy 4083 AAGGCTGATTAAGCAAAACAGTGTGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4142
Db 9280 CTGTTGACACCGCATCAACAGTGTGCTGCTATTAAGAGAGAGAGAGAGAGAGAGAGAG 9339
Qy 4143 TCAATGCGCGAGATCTTATTAAGGACACAGACCACTCAGTGAAGTAAATTTGCACTTC 4202
Db 9340 TCAATCTACCGAGCTTGTATGAGATGAGACACACAGCGCCCAAGGAAATTTGATTTGC 9359
Qy 4203 CATTCGCTTGAACCGACAGTGTGCGCGGTTCCGTTAGCTACACGCTTACAGTACGA 4262
Db 9400 CTTTAAAGTTGATCCGCGGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 9459
Qy 4263 AGTGTTCAAAGGATATCACTTCACTGATGATGATGATGATGATGATGATGATGATGATGAT 4322
Db 9460 ATGGCTTTAAACATCAGCTCCAAATTAAGATCAGACCACTTGAATGCTCACACCA 9519

OY	4423	GAAATTTGGGGCTCGAGCAGACGGCAACAGAGAAATGAAATTAACGGGCTTACATCCAGGA	4382
Db	9520	GGAGACTTAGGGGCAAAACCCGGAAACCAACCACTGATATGATCTGTGGAAAGACGGTAGAA	9579
OY	4383	ATTTTCTGTGGGGCGAAGAGGGCTGAGATACGTATGGGGTAAACATGAAACAGTACAGAG	4442
Db	9580	ACTTACCGGTGACCGAGATGGCTCTGGAAATACATATGGGAAATCATAGGCCAGTAAAGG	9639
OY	4443	TCTGGGCCCAAGGATCGGCACACGAGGCCACCATGTGATGGCCGATGAGATCATATCC	4502
Db	9640	TCTATGCCCAAGATCAGCAACAGAGAACCTTCACGGATGGCCACAGAAATAGTACAGC	9699
OY	4503	ACTATTAATCATCGGATATCAGTGTACACTGTCAATGTGCTGTATGTGTGCTCTTGCTA	4562
Db	9700	ATTACTACATCGCCATCTCTGTGACACATCTTAGCCGTGATAGCTACCGTGGCA	9759
OY	4563	TCTGTGATAGGCACTGTACATCAGACCACTTGCACTCGCCAAAGCAAGAGAGATCGCTGA	4622
Db	9760	TGATGATTTGGGGTAACCTGTGGTACATGTATATGTGCTGTAAAGCGCCGGATGGCTTGA	9819
OY	4623	CGCCATACGGCGCTTGCAACCGAAGCGAACGGTACCCACAGCAATTAGCGGTTTTGTGTGA	4682
Db	9820	CGCCATACGCGCTTGCGCCCAAAACCCCGTAATCCCACTTGCTGACACTTGTGCTGGC	9879
OY	4683	TTGGGCCAACAAAGCTGTGAAACATTTGGAGAAATTTGAACCATCTGTGTTTAAACAC	4742
Db	9880	TTAGGTGGCGCAAAAGCTGAAACGTTTCCAGAGACCAAGATTTACTGTGTGCMAAGTTC	9939
OY	4743	AACGTTTCTCTGGGACAGTTGTGCAATTCCTCTGGAGCGCTGTGTAATCTTTCGCT	4802
Db	9940	AGCGTTCTCTGGGTCCAGTTGTGCACTTTTGGCCGGTTCATCTTCTTAATGCGCT	9999
OY	4803	GCTTTTCACTGTCATGCTCTTTTAAATGTGTGCAAGGCGTCTGCTGGGAGAGTAGAC	4862
Db	10000	ACTGCTCTGCTGCTCTCTCTTTTAAATGTGTGCTGGCGCGCTTACTGGCGAAGTAGACG	10059
OY	4863	CCTTGAAACATGGGACCACTGTGTGCAAAATTTCCGGGATATCCCGTAAAGGCGTGTG	4922
Db	10060	CTTACGAAACATGCGACCACTGTGTCCAAATGTGCCACAGATACCGTAAAGGCACTTGTG	10119
OY	4923	AACCGCAGGTTAAGCGCCACTTAACTGTGAGATCAAGGTGCTCATGCGAATTAACAC	4982
Db	10120	AAAGGCGAGGTAATGCCCCCTCAATTGGAGATCACTGTCAATGTCTCTCGAGGTTTGC	10179
OY	4983	CTTCACTTAACAGAGAGTACGTGACCTGTCAAAATTTCCACACAGTCATTTCTTCCACAAG	5042
Db	10180	CTTCCACCAACCAAGAGTACATTACTGTCAAAATTTACCACTGTGTGCTCCCTCCCAAAA	10239
OY	5043	TTAAATGCTGGGCTCCCTCGAGTGTGCAAGGATCTTCAAAAGCGGATTAACATGCGCGG	5102
Db	10240	TCAATATCTGGGCTCTTGGAAATGTAGCGCGCGCTCATGTGACATATCTCTGCAAG	10299
OY	5103	TTTTTGGGGGTGTACCTTTTATGTGAGGAGGGCGCAATGCTTCTGTGACATGAGA	5162
Db	10300	TCTTTCGAGGGGTCTTACCCCTTTATGTGTGGGAGAGCGCATGTTTTGTGAGATGAGA	10358
OY	5163	ACACACAACACTGATGAGGCGTACGTCAAGTTGCTCCAGACTGTCACTATATGATCAAGCAG	5222
Db	10360	ACACCGAGATGATGAGGCGTACGTCAATTTGTACAGAAATTCGGGTCTGACACAGGCC	10419
OY	5223	TTCGACTTAAAGTTTCAACAAGCTGTCTTGAAGTCGGCTGTGATATGATACGSCACA	5282
Db	10420	AGGCGATTAAAGTGTACACTGCGCGCGATGAAGTAGGACTGCGTAAATGTCTACGGGAA	10479
OY	5283	CCACCGGCGACCTTGATACGTTTGTCAATGGCGTACCGCAGGTTTCTTACCGGACCTGA	5342
Db	10480	CTTACACGTTTCTTGAATGTGTACGTGAACGAGTACACACAGAAAGCTTAAAGACTTGA	10539
OY	5343	AGGCTATAGCAGGCGCATATCAAGCCGCTTTTTCACCTTTGACCAATTAAGTGTGTATCA	5402
Db	10540	AAATCATATGCGAGCAATTTACAGCATCTGTTTACGCATTTGCATCATTAAGTGTGTATCC	10599
OY	5403	GAAAGGGGCTTTTACACTACGACTTCCCTGTAGTATGAGGCTATGAACACAGAGCCT	5462

Dd		10600	AATCGGGCCCTGGTGTACAACTTAAGACTTTCCCGAATAATGAGCATGAACAAGAGCCT	106599
Oy		5463	TGGCGGATATTCAAGCATCTTCGCTTGATGCTAACAGACATAGTAGGCCCGCACTGCATPAC	5522
Dd		10660	TCGGAGACATTTCAAAGTACTCTCTTAGTACAGAAAGATCTCATTCGCACAGACATTA	10719
Oy		5523	GCGTGTGAAGCCTTCTGTCAAGAATCCAAGTCCCCCTAACACCAGACAGATACAGGT	5582
Dd		10720	GCGTACTCAAGCCCTTCGCGCAAAAGTGCATGTCCCCTACACAGCGAGGCCGATCAGAT	10779
Oy		5583	ATGAATATGAGAAACAACTCACAGAGCAAGCCCTGCAGAAAACAGACCAATTGGATGTA	5642
Dd		10780	TTGAGATGTGAAAAACAATCCAGGCCCGCCACTGCAGAAACCGCACTTTTCGGGTGTA	10839
Oy		5643	AAATTTAAGTGAAGCCTCTGCGAGCCTCTTAACCTGTGCTTAACGGGACATCCCTATCTCGA	5702
Dd		10840	AGATTGCAATTAATTCGGCTCGAGCGGTGGACGTGTTACACGGSAACATTTCCATTCTCA	10899
Oy		5703	TTGACATCCCTGATGACAGCTTTTGTGAGATCATCAGATCACCACAAATTTTGAAGTTA	5762
Dd		10900	TTGACATCCCGAAGCGCTGCTTTATCAGACATCAGATCACCACTGTGCTCAACAGTCA	10959
Oy		5763	GCATGCACATGACGACATCGATTTATTCTTGCAGACTTTTGTGTGTTCTCTAACATTAAGT	5822
Dd		10960	AATGTGAAGTCACTGTGATGACTTTATTCAGACAGACTTCGCGGAGATGGCCACCTGCAGT	11019
Oy		5823	ACAAAGCTGACAGAGGAGGACATTTGTCCAGTTCACTCCCACTCCACAGACAGCTGTTTTGA	5882
Dd		11020	ATGATATCCGACCCGGAAGGTGCATGCCCCGTATCTTCGATTCGAGACAGCACTCTCC	11079
Oy		5883	AGGAAGGACACACATGTGACTGCTCCGTAGGCGACATTAACCTACATTTTACACATCGA	5942
Dd		11080	AAGAGTGCAGATACATGTCCTTGAGAAAGAGCGGTGACATGACACTTTACACCGCGA	11139
Oy		5943	GCCCAACAAGCAATTTTATAGTTTTCGCTATGCGGCAAGAAAGTCCACCTGCATGCTGAT	6002
Dd		11140	GTCACACGGGAACTTTATGTAATGCTGTGTGGAAAGAACACATGCAATGACAGAT	11199
Oy		6003	GTAACACACCGGCGGACCAATTAATTGGAGAAACACATAAAGTCGACCAAGAATTCAGG	6062
Dd		11200	GTAACACACACGCTGACATATGTGAGACACCCCGCAAAATATGACCAAGAATTTCAAG	11259
Oy		6063	CGGCAGTTTCCAAACATCTTGGAACTGCTGCTGCTGCACTGTTTGGGGAGCATATCC	6122
Dd		11260	CCGCACTCTCAAAAACATCATGGAAGTTGACTGTTTSCCTTTTCGGCGGCGCTGTCGC	11319
Oy		6123	TCATTTGTTGAGACTTTATAGTGTGTGCTGCAGCTCTATGCTTTAAACAACAGTAAAT	6182
Dd		11320	TATTAATTAATAGACTTAATGATTTTGTCTGACGACATGATGTGATGACACACAAAGT	11379
Oy		6183	GACTGAGCGGAGACATGTA-CATAGCGGTAAACATGATGTACTTCCAGSAAAGGTGAT	6241
Dd		11380	GACCGTACGCCCCCAATGATCGACCAAGCAAACTGATGTACTTCCAGSAACTGATGT	11439
Oy		6242	GCAATATGCGCACGCGCGCT 6261	
Dd		11440	GCAATATGATCATGAGGCTGAT 11459	
RESULT 15				
ID	AAZ92765			
ID	AAZ92765	standard; DNA; 1656 BP.		
XX	AAZ92765;			
XX				
XX	AC			
DT	22-MAY-2000	(first entry)		
DE		Representative eukaryotic layered vector initiation system SEQ ID NO:1.		
KM		Recombinant alphavirus vector; gene therapy; anticancer; antiviral;		
KM		eukaryotic layered vector initiation system; antimicrobial; cardiant;		
KM		antidiabetic; antineurodegeneration; immunomodulatory; immune response;		

OY 1960 AACGGTTAACCACTGAGCGTGGCGGCTCTATCGCGCAGCATATAGTCATGTTG 2019
DB 7123 AGCGGCTTAAACGTCAGATGTGCAGGCTTTCATTTGGGAGCAGCAACATCATACATGAG 7182
OY 2020 TCGTCTCCGACACCTTTGATGCGGAGATGCGCCACTTGGCTGAACATGAGTAATA 2079
DB 7183 TAGTATCTGACAAAGAAATGGCTGAGAGGTGCGCACTGGCTCAACATGAGGTTAAGA 7242
OY 2080 TTATGATGCACTTATTTGGTATCAAGCACTCTCTGTGGGGATTTATCCGTGG 2139
DB 7243 TCATGACGCGATCATCGGTGAGAACCACTTACTCTGCGCGGATTTATCTTGCAG 7302
OY 2140 ACCAGATTAACAGGCACTGCTGAGATGCGACAGCTCTTAAAGAGCTTTTAAAGCTTG 2199
DB 7303 ATTGGTTACTTCCACAGGCTGCGGTGGGATTCCTCTGAAGAGCTGTTTAAAGTTGG 7362
OY 2200 GAAACCACTTGGCACTCGATATACCAGACTGCGACCGCGCGGCACTGCATGATG 2259
DB 7363 GTAAACCGCTCCAGCCGACGAGCAGAGCAAGACGAAGACGCGCTGTGATGATG 7422
OY 2260 AAGCAATGCAATGAAACGAATTTAGGAGAGTATGAAAGCGCGTGAAGATCCA 2319
DB 7423 AAACAAAGCGGTGTTTAAAGTATTAACAGGCACTTACAGTGGCGTGAAGAGCC 7482
OY 2320 GATACGAGATCATACCTGCGAGGCTGATCATACGCTCTCTGCAAGTTAGCGGAAGCG 2379
DB 7483 GGTATGAGTGAACAATTTACACCTGTCTACTGGCATTTGAGAACTTTGGCCAGAGCA 7542
OY 2380 TTAAAGACTTCAGAGCATAGAGGAGCCCAATCACTCTTACGCGTGAACCTTAATAG 2439
DB 7543 AAAGAGCAATCCAAACCATCAGAGGGGAAATTAACATCTTACGATGCTCTTAATAGT 7602
OY 2440 TGAGCTATGACAGGCACTTAC-----CCACCGGCGAAGATTTCCATACC 2486
DB 7603 CAGCATAGTACATTTCACTGTGACTAATCTACACACCAACCACTGATAGAGATCT 7662
OY 2487 CTCAGCTGAATCTTCCACCAAGTTTACCTTACCAATCCGATGGCTTACCGAGATCCAAAC 2546
DB 7663 TTAACATCTCGGCGCGCGCCCTTCCGCGCCCACTGCCATGTGAGAGCGCGGAGAA 7722
OY 2547 CTCTAGGCGCGCTGAGAGCGGCTTTCGCGCCCGCTGCTGCTCAATTCGAAGATCTTA 2606
DB 7723 GGAGCAGAGCGGCGCCGATGTCTGCGCCCAAGCGGCTTCTCAAACTCCAGCACTGA 7782
OY 2607 GGAAGTCAATGATCACTTGAATTTCAACCAACGATCACCTTATCCGCCCGCAGGTCCA- 2665
DB 7783 CCACAGCGGTACAGTCCCTTATGATGACAGGCAACTAGACCTCAACCCCGACGTCAC 7842
OY 2666 -----CCGCCAAAGAAAGAGTGTCTCTTACGCCAAACCTTACTGAGCTTAAAGA 2720
DB 7843 GCCCGCCACCGCGCAGAAAGAGCAGGCGCCCAACCAACCGAAGCCGAAAGAAACCA 7902
OY 2721 AGAAGCAGCAAGCCAAAGAGAGCAAAAC--GCAGCTTAAACAGGAGAAACGACAACTGA 2777
DB 7903 AAACCGCAGAGAAAGAAAGAAAGCAACTGCAAAACCCAAACCCGAAAGAGACGCCA 7962
OY 2778 TGTGTATGAAGTTGAGTTCGAGCAGACATTTCCGATCATG--CTGAAGCGCCAGTGA 2834
DB 7963 TGGCATTAAGTTGAGGCGCGCAGATTGTTGACGTCAGAGAACAGAGACGAGATGTCA 8022
OY 2835 ATGATATAGCTGCTGTTCGAGAGAGAGCTGATGAACCACTCCAGCTTGAAGGAAAGA 2894
DB 8023 TCGGCGACGCACTGCGCATGAGAAAGATTAATTAACCTTGCACGTGAAGAGAACCA 8082
OY 2895 TTGATATAGCAATTAGCGCGCGGTGAATTTGAAGAGGCTAGCATGTATGCACTTGAAGT 2954
DB 8083 TCGACCACTCTGTGTATCAAAAGCTCAATTTAACAAAGTGTCAACATATGCAATGAGT 8142
OY 2955 ACGGAGAGCTTCCCGAAGCATGAATTCAGACAGCTGACATGACACAGCGCAAAACCA 3014
DB 8143 TCGACAGTGTGCAAGTCAACATGAGAGGCACTTCACTTACACAGTGAACACCCCG 8202

OY 3015 CGGCTTCTACAACTGAGCAGCACGCGAGCTCCAGTATAGAAATGAGATTTTACCGTAC 3074
DB 8203 AAGATTTCTATATCTGGCAGCACCGAGGGGTGCACTTATAGTGAGATGATTTACATCC 8262
OY 3075 CGAGAGAGTGGCGGGAAGAGCGACAGCGGAAGACCGATCTGTGACAAACAGAGCAGAG 3134
DB 8263 CTCGCGAGTGAAGAGCAGAGAGACAGCGGCTGTCCATCATGTATTAATCCGGTCCGG 8322
OY 3145 TTGCGCATATGTTCTAGAGAGTGAATTAAGAGGAGACCGTACCGGCTTTGAGTGTCA 3194
DB 8323 TTGTGCAATAGTCTTGGTGGCTGTGATGAAGGAACAGAACTCCCTTGTGCTGCA 8382
OY 3195 CTTGGAACAGAAAGGGGTGACCATTAAGAGATACCCCGAAGTTCTGAACCGTGT--- 3251
DB 8383 CTTGGAATAGTAAAGGAAGAACAAATTAAGAGAGCCCGGAAGGAGACGAAGAGTGTCCG 8442
OY 3252 -----CACTAGTTACAGCGCTATGCGTCTTGAATGTACGTTCCATCGCAAAAC 3305
DB 8443 CAGCACACTGGTACCGGCAATGTGTTGCTCGAAATGTGAGCTTCCATCGACCGCC 8502
OY 3306 CACCGGTGCTATTCATGACGCGAAGCAAGCACTGAGCGTGTGGAAGGAAGCTGCG 3365
DB 8503 CGCCCATGTCTATACCGCGCAACTTCAGAGCCCTGACATCTTGAAGGAAGCTGA 8562
OY 3366 ACAATCCAAATTTACGACACGCTGCTGAGAAAGTCTTGAATGTCCA--TCACCGCGGC 3422
DB 8563 ACCATGAGGCGCTAGAGTACCTGTCTCAATGCTATTTCCGGTGGGATGTGTGGCAGAA 8622
OY 3423 CCAAACGAAGCATTAACGATGACTTACACTGACAGTGTCTTACTGCGGTTCTGCGCT 3482
DB 8623 GCAAAAGAGGCTGCTTACGACATTTTACCTGACAGCGCCCTTACTTGGGCACTCTCGT 8682
OY 3483 APTGAGACACTCAACGCGGTGTTTCAAGCCCAATTAATAATGAGAACGTGCGGAGAT 3542
DB 8683 ACTGCACATATCTGAACCGTGTCTTACCGCTTGAATGAGAGAGCTGTGGACGAAG 8742
OY 3543 CTGATGATGATTCGATTAATGAAATCCAGGTCTGCGGCAATTTGGCTTACATCAGCAGCA 3602
DB 8743 CGGAGATTAACCATATGAGTACAGACTTCCGCCAATTTGAGATACGACAAACCGAG 8802
OY 3603 CTGCGGATGTACCAATTTCCGTTACATGTCTTTGACCAAGACCATATCAAGAGAG 3662
DB 8803 CAGCAAGCGCAAAACGATACCGCTTACATGTGCTTAAACAGGATCACACCGTTAAAGAG 8862
OY 3663 ACAATGAGAGAAATTAATCTATCAGCAGACTGTGACCTGCGCTGTCTTGGCCCAAG 3722
DB 8863 GCACCATGATGACATCAAGATTAAGACCTCAGAGCGGTGAGAGGCTTAGCTTCAAG 8922
OY 3723 GGTACTTCTGTAGCTCAATGTCTCTCAGGTTGACAGTGAACGCTCAGTATCACAGAG 3782
DB 8923 GATACTTCTCTCGCAAAATGCTCTCCAGGGGACAGGTTAAGGTTAGCATAGTAGTA 8982
OY 3783 GAGCATCTGAGAAATTCATGACCGGTGAGAAAGATCAGAGAGAGTTGTGCTAGAG 3842
DB 8983 GCAACTCAGCAACGTCATGTACACTGCGCCCAAGATTAACCAAAATTCGCGGAGCGG 9042
OY 3843 AGAGTACTTGTCCACCGGCTCATGAGAAAGCTGTAAAGTCCACTTTACGATCACT 3902
DB 9043 AAAAATATGATCTACTCTCCGTTCAGGTAAAGAAATTTCTTGCACAGTATGAGCCGTC 9102
OY 3903 TGAAGAGAGCTGCGCGGTATACATTAACATGACAGGCGCAGGCGCAACGCGTATAGT 3962
DB 9103 TGAAG--ACAACGTGAGGCTATCATCTATGACAGGCGCGGACCGCATTTATCACT 9159
OY 3963 CTTATCTGAGAGAGGCTCAGGCGAAGTGTACATTTAAACACTTCTGCGCAAGAGCTCA 4022
DB 9160 CTTACTCTGAGAAATCATCAGGAGAAAGTTTACGCAAAACCCGCACTCTGGAAGAACATTA 9219
OY 4023 CTTAGCAATTTAAAGTGTGCGCATACAGCAAGAGTATGTTGAGACGCGGAACGAAGATGA 4082
DB 9220 GGTATGATGTGAAGTGTGCGGCACTTACAGAGCCGGAACGTTTTCGACCGGCAACGAATCA 9279
OY 4083 ACGGCTGACTTAAAGCAAAACAGTGCATTTGCTTACAGAGGAGCAACGAATGCTCT 4142

Db 9280 CTGGTTGACCCGCATAGAGAGTGCCTCGCTTAATAGAGGACCAACGAAGGGGCT 9339
Qy 4143 TCAACTCGCCGATCTTAATAGGACACAGACCACTAGTCGAAGTAATTGCACTTC 4202
Db 9340 TCAACTCACCGGCTTATGATGACATGACGACACACGGCCCAAGGAAATTTGATTTGC 9339
Qy 4203 CATTCGCTTGACACCGACAGTGTGCGGTTCCGTTAGTGCACACGCTTACAGTCA 4282
Db 9400 CTTTCAAGTTGATCCCGGTCCTGCAATGCTCTGTTGACACGCGCCGAAATTAATAC 9459
Qy 4263 AGTGGTTCAAGGATCACTCCACCTGACTGCAATGCGACCAATGCTGCAACGA 4322
Db 9460 ATGGCTTAAACATCAGGCTCCAAATTAGTACAGACCACTTGAATGTGTCCACCA 9519
Qy 4323 GAAATTTGGGCTGCGACGACGCAACAGCAATGATTAACGGGCTTACATCCAGA 4382
Db 9520 GGAGACTAGGGGGCAACCCCGAACCAACTGAATGATCTGTGAAAGACGGTCA 9579
Qy 4383 ATTTTCTGTGGGGCGAAGGGCTGAGTACGTATGGGGTAAACATGAACAGTCAAG 4442
Db 9580 ACTTACCGTCGACCGAGATGCTGGAATACATATGAGGAATCATAGACAGTGAAGG 9639
Qy 4443 TCTGGGCCAGAGTCCGACACGACGACCAATGATGCGCATGAGATCATATCC 4502
Db 9640 TCTATGCCAAGATGACGACGACGAGACCTCTCAGATGCGCACAGAAATATGTAAGC 9639
Qy 4503 ACTATTATCATCGGATCTCAGTCTTACACTGTCAATTGTGTGTGTGTGTGTGTGT 4562
Db 9700 ATTACTACATCGCATCTGTGTGACCACTTATGAGCGGTGCGATCAGTCACTGTGCGA 9759
Qy 4563 TCCGTGAGGCACTGATCATCAGCACTTGATGCGCAAGCAAGAAAGACGCTGCA 4622
Db 9760 TGAATGATGCGTAACTGTGCACTGTGCACTGTGCACTGTGCACTGTGCACTGTGCA 9819
Qy 4623 CGGCATACGCGCTTGACACGACCAACGATACCCAGACATTAAGCGGTTTGTGCTGA 4682
Db 9820 CGCATATGCGCTTGCGCCCAAGCGCGTAATCCCAATCTGCGTGGCACTTGTGCTGCG 9879
Qy 4683 TTGGGCCAACCAACGCTGAAACATTTGAGAACTTTGAACATCTGTGTTTAAACAAC 4742
Db 9880 TTAGTGTGCGCAATGTGAAACGTTCAACGAGACCATGATTAATCTGTGCGAACATC 9939
Qy 4743 AACGTTCTGTGGGACAGTGTGCACTCTCTGAGAGCGCTTATCTGTGCTGCT 4802
Db 9940 AGCGTTCTGTGGGACAGTGTGCACTCTCTGAGAGCGCTTATCTGTGCTGCTGCT 9999
Qy 4803 GCTTTTCATGCTGATCCTTTTATTTGTTGTTGAGGCGTCTGCTGAGGAGGTAGACG 4862
Db 10000 ACTGCTCTGCTGCTGCTTTTATTTAGTTGCTGCGGCGCTTACCTGCGAAGGTAGACG 10059
Qy 4863 CTTTGCACATGCGACATCTGTGCAAAATGTTCCGGGATCCGTAATAGGCGTTGTGCTG 4922
Db 10060 CTTACGAAATGCGACATCTGTGCAAAATGTTCCGGGATCCGTAATAGGCGCTTGTGCTG 10119
Qy 4923 AAGCGGAGGTTAGCGGCACTTAACGAGATCAGGTCGTCAATCGGAATTAACAC 4982
Db 10120 AAGGGAGGAGTATGCCCGCTCAATTTGAGATCATGTATCTCTCGAGGTTTTC 10179
Qy 4983 CTTTCAATCAAGAGATGATGACCTGCAAAATTCACACAGTCAATCTTTCACACAG 5042
Db 10180 CTTTCAACCAAGAGATGATGACCTGCAAAATTCACACAGTGTGCTCTCTCCCAAAA 10239
Qy 5043 TTAATGCTGCGGCTCTGAGTGAAGGATCTTAAGGCGGATTAACATGCGCG 5102
Db 10240 TCAATGCTGCGGCTCTGAGTGAAGTGAAGGCGGCTTCAATGACATTAATCTGCAAGG 10299
Qy 5103 TTTTGGCGGTGTACTCTTTTCAATGTGGGAGGCGCAATGCTTGTGACAGTGA 5162
Db 10300 TCTTGGAGGGGTCTACCCCTTAATGTGGGAGGAGCAATGTTTGTGCAAGTGA 10359
Qy 5163 ACACACAATGATGAGGCGTACGTGAGTTGCTTCCAGATGCAATATAGATACCGAG 5222

Db 10360 ACAGCAGATGATGAGGCGGTACGTGATTTGTACAGAGATTGCGCTGTGACACCGCG 10419
Qy 5223 TCGCATTAAGTTTCAACAGCTGTCTTGAAGTCCGCTCGTATATGTTATGCGCAACA 5282
Db 10420 AGCGATTAAGGTGACACAGTCCGCGAGTAAGATGAGTGTATATGTATGCGGAA 10479
Qy 5283 CCAACCGGACCTGATACCTTTGTCAATGAGCTTCAAGCGGATCTCTCAAGGACCTGA 5342
Db 10480 CTACAGATTTCTAGATGTGTAGTGAAGAGATGACACAGGAACCTTAAGACTTGA 10539
Qy 5343 AGGTATAGAGGCGCGATATCAGCGCTTTTCACTTTGACATATAGGTGTATCA 5402
Db 10540 AAGTATAGCTGAGACCAATTTCAAGATCTGTTACGCTATTCATTAAGTGTGTATTC 10599
Qy 5403 GAAAGGCTTTGTTTCAATGACCTTCCGTAGTATGAGATGAGAAACCGAGAGCT 5462
Db 10600 ATCGGCGCTGTGTTTCAATGATGACTTCCGAAATATGAGAGATTAACCGAGAGCT 10659
Qy 5463 TCGCGATATTTCAAGCATCTCTCGCTGATGCTACAGATAGTACCGGACATGACATAC 5522
Db 10660 TCGAGACATTTCAAGCTACTCTCTTGAATGACAGATCTCATGCGACGACAGACTTA 10719
Qy 5523 GCGTCTGAAGCTTCTGTCAAGAACATCAAGTCCCTTACACCCAGCATATCAGGT 5582
Db 10720 GCGTACTCAAGCTTCCGCAAGAACGTGATGCTCCGTACAGCGAGCGCATACAGAT 10779
Qy 5583 ATGAAATGTGAAAGAACATCTCAGACGACCCCTGCAAGAAACGACCAATTTGATGTA 5642
Db 10780 TTGAGATGTGAAAGAACATCTCAGACGACCCCTGCAAGAAACGACCAATTTGATGTA 10839
Qy 5643 AAATGAGGTGAGGCTCTGCGAGGCTTAACTGTGCTTACGGGACATCTTATCTGA 5702
Db 10840 AGATGAGTGAATTCGCTCCGAGCGGTGAGCTGTTCATACGGAATCTTCCATTTCTA 10899
Qy 5703 TTGACATCTCTGATGACGCTTTTGTAGATCATAGATCAACCAATTTTGAAGTTA 5762
Db 10900 TTGACATCTCTGATGACGCTTTTGTAGATCATAGATCAACCAATTTTGAAGTTA 10959
Qy 5763 GCTGACAGTGAAGTGAATTTTGTAGATCATAGATCAACCAATTTTGAAGTTA 5822
Db 10960 AATGAGTGAAGTGAATTTTGTAGATCATAGATCAACCAATTTTGAAGTTA 11019
Qy 5823 ACAAGCTGACAGGAGGACATTTGTCAGATTCATCTCCATCAGACAGTGTTTGA 5882
Db 11020 ATGATTCGACCGCGAGTCAATGCTCCGATCTGATCTGATGAGACAGCAACTCTCC 11079
Qy 5883 AGAAGGACCAACATGATGATGCTGAGGACAGATTAACATTAATTTAGACATCGA 5942
Db 11080 AAGAGTGAAGTGAATTTTGTAGATCATAGATCAACCAATTTTGAAGTTA 11139
Qy 5943 GCCCACAAGAAATTTTATGTTTGTGATGCGGCAAGATGCACTGCAATGCTGAAT 6002
Db 11140 GTCCACAGGAACTTTATGTTTGTGATGCTGATGAGGAAAGAACATGCAATGCAAGT 11199
Qy 6003 GTAACACAGGCGGACCAATATTTGAGAACCAATTAAGTGAACCAAGAAATTCAGG 6062
Db 11200 GTAAACACAGGCGGACCAATATTTGAGAACCAATTAAGTGAACCAAGAAATTCAGG 11259
Qy 6063 CGGCAATTTCAAAACATCTTGAAGTGTGCTGCTGATGCTGTTGGGAGATCATCC 6122
Db 11260 CCGCATCTCAAAAACATCATGAGTGTGCTGTTGCTTTCGCGGCGCTCTGCTGCG 11319
Qy 6123 TCAATGTTGAGACTTAATGTTGTGTGAGCTTATGCTTATTAACACAGTAT 6182
Db 11320 TATTAATTAATGAGCTTAATGTTGTGTGAGCTTATGCTTATTAACACAGTAT 11379
Qy 6183 GACTGACGCGGACACTGA-CATAGCGTAAACTCGATGATCTTCCAGAAAGCTGT 6241
Db 11380 GACGCTAGCGCCCAATATTCGACAGCAAACTGATATCTTCCGAGAACTGAT 11439
Qy 6242 GCATATGCAAGCGCGCT 6261
Db 11440 GCATATGCAAGCGCTGT 11459

Mon Nov 17 10:39:24 2003

us-10-023-649-1_copy_5000_11484.rng

Page 49

Search completed: November 15, 2003, 23:18:21
Job time : 1555 secs

Db 5060 CGGGAATTTCTAGAGCAACCAAGTCGGGATCACTGCTATGCTTGGGGGGCTCTGTTCT 5119
Qy 121 CTAGCTCTGCCACACGCTCGTCAACCGCTGAGCGACTATGACAGCGGCTTTGTGTTA 180
Db 5120 CTAGCTCTGCCACACGCTCGTCAACCGCTGAGCGACTATGACAGCGGCTTTGTGTTA 5179
Qy 181 CAGTGTATGTGATCAAGCGCAACGCTCTAGCTGAGCATCCCTAGTGTCTTGGCTTGG 240
Db 5180 CAGTGTATGTGATCAAGCGCAACGCTCTAGCTGAGCATCCCTAGTGTCTTGGCTTGG 5239
Qy 241 ACCTGACAGCTGCTTCTGACGTCACTGATTCGACCTGAGATTTCAAAGTGCATCAAGCT 300
Db 5240 ACCTGACAGCTGCTTCTGACGTCACTGATTCGACCTGAGATTTCAAAGTGCATCAAGCT 5239
Qy 301 TTGAAGTGAAGAACCATCTGTACAGAGCTTAACCTGCGAGGTGCGAAGCTCTGTGAC 360
Db 5300 TTGAAGTGAAGAACCATCTGTACAGAGCTTAACCTGCGAGGTGCGAAGCTCTGTGAC 5359
Qy 361 TTGGCGAAATATGCAAGACTTCAATCTGCTCTTCCAGTTCTTTCGGACTCAAGAC 420
Db 5360 TTGGCGAAATATGCAAGACTTCAATCTGCTCTTCCAGTTCTTTCGGACTCAAGAC 5419
Qy 421 CAGTACCGGCAACCAAGGAGCGCCCATCCCATCACTAGATCGACGCTTCCGACCTC 480
Db 5420 CAGTACCGGCAACCAAGGAGCGCCCATCCCATCACTAGATCGACGCTTCCGACCTC 5479
Qy 481 CAGTTCCAAAGCCACGAGAGCTAAGTACCAACAACCAAGAGTGTCTGAGAGCATCT 540
Db 5480 CAGTTCCAAAGCCACGAGAGCTAAGTACCAACAACCAAGAGTGTCTGAGAGCATCT 5539
Qy 541 CAGAAGCGAGTATGAGAGAGTACATCCGTCAACACTCCAACTGAAGGATGAGCGGAG 600
Db 5540 CAGAAGCGAGTATGAGAGAGTACATCCGTCAACACTCCAACTGAAGGATGAGCGGAG 5539
Qy 601 CGATATTTTCTCATCCGAAAACAGGCCAAGTCACTTTCACAGAAATCACTAGTCAAT 660
Db 5600 CGATATTTTCTCATCCGAAAACAGGCCAAGTCACTTTCACAGAAATCACTAGTCAAT 5659
Qy 661 GTAAATCTACAAAGACCTTATTTGGATCGGGCCGTCTCAGAGAGTATTAACCCCGGCC 720
Db 5660 GTAAATCTACAAAGACCTTATTTGGATCGGGCCGTCTCAGAGAGTATTAACCCCGGCC 5719
Qy 721 TCGATCTCGAAGAGAGAAAATGTTACAGAGAACTGCAATTTAGCGCTCTGAGGAA 780
Db 5720 TCGATCTCGAAGAGAGAAAATGTTACAGAGAACTGCAATTTAGCGCTCTGAGGAA 5779
Qy 781 ATAGAAGCAGGTATCAATCAAGAAAAGTGAATAATGMAAGCAATTAACAGCGAGCAC 840
Db 5780 ATAGAAGCAGGTATCAATCAAGAAAAGTGAATAATGMAAGCAATTAACAGCGAGCAC 5839
Qy 841 TCATTTCTGGATTTGGGCAATATCTATCATCAAGAGTGAATCTGTGAGTGTAAACAGAG 900
Db 5840 TCATTTCTGGATTTGGGCAATATCTATCATCAAGAGTGAATCTGTGAGTGTAAACAGAG 5899
Qy 901 TCAATTTATCCGTACCAATCTACTCGTCAAGGTAAATTAACAGTTTACATCTGAGAGG 960
Db 5900 TCAATTTATCCGTACCAATCTACTCGTCAAGGTAAATTAACAGTTTACATCTGAGAGG 5959
Qy 961 TCGCGGTTAAACGTGCAACTTAACTTATCAAGAAATTAACCTTACAGTACCAAGTTAT 1020
Db 5960 TCGCGGTTAAACGTGCAACTTAACTTATCAAGAAATTAACCTTACAGTACCAAGTTAT 6019
Qy 1021 GTATTAACAGATGAATCAAGTGTATCTTGAACATGTGTGAGCGGCGCATCGTGTCTAG 1080
Db 6020 GTATTAACAGATGAATCAAGTGTATCTTGAACATGTGTGAGCGGCGCATCGTGTCTAG 6079
Qy 1081 ATACAGCCACTTTTGTCCGGCTAACTGAGAGTCAACCAAGAGATGCTATTTGC 1140
Db 6080 ATACAGCCACTTTTGTCCGGCTAACTGAGAGTCAACCAAGAGATGCTATTTGC 6139
Qy 1141 AGCAGAGATTAAGATCAAGCCGCTCCATCGCTTATCAAGATTAATTAATGATTTGG 1200

Db 6140 AGCAGAGATTAAGATCAAGCCGCTCCATCGCTTATACAGATTAATTAATGATTTGG 6199
Qy 1201 CTGACGTACTTAAGAAATTTGCAAGTTAACCAATGAGAAATTAACCTGCTTATGATT 1260
Db 6200 CTGACGTACTTAAGAAATTTGCAAGTTAACCAATGAGAAATTAACCTGCTTATGATT 6259
Qy 1261 CGCGGCAATTAATGTTGATTTGTTTCAAGAAATATACATGATGATGATCTGGGATA 1320
Db 6260 CGCGGCAATTAATGTTGATTTGTTTCAAGAAATATACATGATGATGATCTGGGATA 6319
Qy 1321 CTTTGGCGATTAACCTTATTCGGCTTAATCAAGAGAACTTACGCAATATGTGACAAAG 1380
Db 6320 CTTTGGCGATTAACCTTATTCGGCTTAATCAAGAGAACTTACGCAATATGTGACAAAG 6379
Qy 1381 TGAAGGGCCGAAGAGAGAGAGCATTTGTTGCGAATCTCATTAATCTTAAACGTTGAGG 1440
Db 6380 TGAAGGGCCGAAGAGAGAGAGCATTTGTTGCGAATCTCATTAATCTTAAACGTTGAGG 6439
Qy 1441 AGATTAACCAATGATCAATTTGCTATGATGATCTTAAAGAGAGATGTCAAAGTTACTCCGGCA 1500
Db 6440 AGATTAACCAATGATCAATTTGCTATGATGATCTTAAAGAGAGATGTCAAAGTTACTCCGGCA 6499
Qy 1501 CGAAACATACAGAGAGAGCGGCTTAAGGTGAGTTATTCAGGCTGCAAGATCCCTTGTCTA 1560
Db 6500 CGAAACATACAGAGAGAGCGGCTTAAGGTGAGTTATTCAGGCTGCAAGATCCCTTGTCTA 6559
Qy 1561 CCGCTTAACCTTTCGGGATTCATCGGGAATTAATCCGTGAGCTGAGATGCGGCTTCTGC 1620
Db 6560 CCGCTTAACCTTTCGGGATTCATCGGGAATTAATCCGTGAGCTGAGATGCGGCTTCTGC 6619
Qy 1621 CAAATATCACTACTCTCTTTCGACATGTCAGCGGAGATTTTGAATGCGATTTGCTGAAC 1680
Db 6620 CAAATATCACTACTCTCTTTCGACATGTCAGCGGAGATTTTGAATGCGATTTGCTGAAC 6679
Qy 1681 ATTTCCACACGCGGACCCAGTATTTGGAACGACATCCGCTGTTGATTAAGGAGAG 1740
Db 6680 ATTTCCACACGCGGACCCAGTATTTGGAACGACATCCGCTGTTGATTAAGGAGAG 6739
Qy 1741 ACGACGCTATCGCATTTGGCGGTGATGATCTTGAAGGACTTAAGGTGAGCAACCGC 1800
Db 6740 ACGACGCTATCGCATTTGGCGGTGATGATCTTGAAGGACTTAAGGTGAGCAACCGC 6799
Qy 1801 TCTTAGATTTGATAGAGCGCGCTTCGGCAATATCACTGTCGCACTTACTTACAGGAA 1860
Db 6800 TCTTAGATTTGATAGAGCGCGCTTCGGCAATATCACTGTCGCACTTACTTACAGGAA 6859
Qy 1861 CGAGTTTAAATTTGGTGCATGATGAATCCGATATGTTCTTAACGCTGTTGTCAACA 1920
Db 6860 CGAGTTTAAATTTGGTGCATGATGAATCCGATATGTTCTTAACGCTGTTGTCAACA 6919
Qy 1921 CACTAGCAATATCAATGATTTGCTAGAGAGTACTAGCTGGAACGGTTAACACGTCAGCGT 1980
Db 6920 CACTAGCAATATCAATGATTTGCTAGAGAGTACTAGCTGGAACGGTTAACACGTCAGCGT 6979
Qy 1981 GCGGCGCTCTATCGGCGAGATTAACATAGTGAATGCTGTCTCGACACCTTGAATG 2040
Db 6980 GCGGCGCTCTATCGGCGAGATTAACATAGTGAATGCTGTCTCGACACCTTGAATG 7039
Qy 2041 CGAGAGATCGGCACTTGGCTGAACATGGAAGTAAATTAATGATGCAATTTGGTA 2100
Db 7040 CGAGAGATCGGCACTTGGCTGAACATGGAAGTAAATTAATGATGCAATTTGGTA 7099
Qy 2101 TCAAGACACCTTCTTGTGGGGGATTTATCTGCTGAGACAGATTAACAGGCAACGCT 2160
Db 7100 TCAAGACACCTTCTTGTGGGGGATTTATCTGCTGAGACAGATTAACAGGCAACGCT 7159
Qy 2161 GCAGAGTCGAGACCCCTTAATAAGGCTTTTAACTTGAAGAAACATTTGCCAGTGCATG 2220
Db 7160 GCAGAGTCGAGACCCCTTAATAAGGCTTTTAACTTGAAGAAACATTTGCCAGTGCATG 7219
Qy 2221 ATACCAAGACTGAGACCGGCTCGGCGACTGATGATGAAGCATGCGATGGAACAGAA 2280
Db 7220 ATACCAAGACTGAGACCGGCTCGGCGACTGATGATGAAGCATGCGATGGAACAGAA 7279

QY 2281 TTGGAATTAACGACGAGTATGAAAGCCGTAGAAATCCAGATACGAGATCATACTGCGAG 2340
Db 7280 TTGGAATTAACGACGAGTATGAAAGCCGTAGAAATCCAGATACGAGATCATACTGCGAG 7339
QY 2341 GCTTATCATCATCGTCTCTGTCGACGTTAGCCGAAAGCGTTAAGAACTTCAAGAGCATTA 2400
Db 7340 GCTTATCATCATCGTCTCTGTCGACGTTAGCCGAAAGCGTTAAGAACTTCAAGAGCATTA 7399
QY 2401 GAGGAGGCCATCAACCCCTTACGCGTACCTTAATAGGTGACGTATGAGAACCCACCT 2460
Db 7400 GAGGAGGCCATCAACCCCTTACGCGTACCTTAATAGGTGACGTATGAGAACCCACCT 7459
QY 2461 ACCCAACCGGACAGAAATGTTTCCATACCTTCAGCTGAACTTTCCACAGTTTACCTTACAA 2520
Db 7460 ACCCAACCGGACAGAAATGTTTCCATACCTTCAGCTGAACTTTCCACAGTTTACCTTACAA 7519
QY 2521 TCCGATGGCTTACCGAGATCCAAACCTCTTAGGCGCGCTGAGAGCGCTTTCGGCCCC 2580
Db 7520 TCCGATGGCTTACCGAGATCCAAACCTCTTAGGCGCGCTGAGAGCGCTTTCGGCCCC 7579
QY 2581 GCTGCTGCTCAATGSAAGATCTTAGAGAGTGCATAGTCACTTGAACCTTCAACCAACG 2640
Db 7580 GCTGCTGCTCAATGSAAGATCTTAGAGAGTGCATAGTCACTTGAACCTTCAACCAACG 7639
QY 2641 ATCACTAATCCGCGCCAGTCCACCGGCAAGAAAGAGAGTGTCTTAAGCAAA 2700
Db 7640 ATCACTAATCCGCGCCAGTCCACCGGCAAGAAAGAGAGTGTCTTAAGCAAA 7699
QY 2701 ACTTACCTAATCCGCGCCAGTCCACCGGCAAGAAAGAGAGTGTCTTAAGCAAA 2760
Db 7700 ACTTACCTAATCCGCGCCAGTCCACCGGCAAGAAAGAGAGTGTCTTAAGCAAA 7759
QY 2761 AGGGAAGAGACACGTATGTATGATAGTGAAGTTCGACACAGACATTTCCGATCAATGT 2820
Db 7760 AGGGAAGAGACACGTATGTATGATAGTGAAGTTCGACACAGACATTTCCGATCAATGT 7819
QY 2821 GAAACGGCAAGTGAATGATATGCTGCTGCTGCGAGGAGAGGCTGATGAACCACTCA 2880
Db 7820 GAAACGGCAAGTGAATGATATGCTGCTGCTGCGAGGAGAGGCTGATGAACCACTCA 7879
QY 2881 CGTTGAAGAAAAATTTGATATGACCAATTAAGCGGCGGTGAATTAAGAAAGCTAGCAT 2940
Db 7880 CGTTGAAGAAAAATTTGATATGACCAATTAAGCGGCGGTGAATTAAGAAAGCTAGCAT 7939
QY 2941 GTACACATTTGAGATACGCGGAGCTTCCCGAAGACATGAATACAGACGCTGCAATCAC 3000
Db 7940 GTACACATTTGAGATACGCGGAGCTTCCCGAAGACATGAATACAGACGCTGCAATCAC 7999
QY 3001 CAGCGACAAACCAACCGGCGCTTCTACAACTGGGACCAACCGGCGGATGAGAAATG 3060
Db 8000 CAGCGACAAACCAACCGGCGCTTCTACAACTGGGACCAACCGGCGGATGAGAAATG 8059
QY 3061 GAGATTTACCGTACCGAGAGAGTGGCGGAAAGCGGACAGCGGAAAGCCGATCTGGA 3120
Db 8060 GAGATTTACCGTACCGAGAGAGTGGCGGAAAGCGGACAGCGGAAAGCCGATCTGGA 8119
QY 3121 CAACAGAGGACAGATTTGAGCTATTTGTTAGAGAGTGAATGAGGGACGCGTACGCG 3180
Db 8120 CAACAGAGGACAGATTTGAGCTATTTGTTAGAGAGTGAATGAGGGACGCGTACGCG 8179
QY 3181 GCTTCAAGTGTCACTTGGACCAAGAAAGGGTGAACATTAAGGATACCCCGAAAGGTTG 3240
Db 8180 GCTTCAAGTGTCACTTGGACCAAGAAAGGGTGAACATTAAGGATACCCCGAAAGGTTG 8239
QY 3241 TGAACCGTGTCTAGTATACAGCGTATGCGTCTTTCGAAATGTCAAGCTTCCATGCGA 3300
Db 8240 TGAACCGTGTCTAGTATACAGCGTATGCGTCTTTCGAAATGTCAAGCTTCCATGCGA 8299
QY 3301 CAAACCAACCGTGTCTATTCATGACGCGCAAGACAGACATGACGCTGGAAGAGAA 3360
Db 8300 CAAACCAACCGTGTCTATTCATGACGCGCAAGACAGACATGACGCTGGAAGAGAA 8359

QY 3361 CGTGCAATCCAAATTACGACACGCTGAGAGACGTCTTAATGTCATCACGCGG 3420
Db 8360 CGTGCAATCCAAATTACGACACGCTGAGAGACGTCTTAATGTCATCACGCGG 8419
QY 3421 GCCCAACGAAGCATTTACCGATGACTTCACTGACCACTGCTTACCTGGGGTTCTGCC 3480
Db 8420 GCCCAACGAAGCATTTACCGATGACTTCACTGACCACTGCTTACCTGGGGTTCTGCC 8479
QY 3481 GTATTGACAGACCTCAAGCGGCTGTTTACGCGCAATTAATTAAGAACTGTGGAGCA 3540
Db 8480 GTATTGACAGACCTCAAGCGGCTGTTTACGCGCAATTAATTAAGAACTGTGGAGCA 8539
QY 3541 ATCTGATGATGATGATTAAGATTCAGATCCAGTCTCGGCAATTCGCTCAATCAGGCA 3600
Db 8540 ATCTGATGATGATGATTAAGATTCAGATCCAGTCTCGGCAATTCGCTCAATCAGGCA 8599
QY 3601 CACTGCGGATGTCACCAAAATTCGTTACATGCTTTTCCAGACGACATGACATCAAGGA 3660
Db 8600 CACTGCGGATGTCACCAAAATTCGTTACATGCTTTTCCAGACGACATGACATCAAGGA 8659
QY 3661 AGACGATGAGAAATTAAGATTCAGACATCTGACCCCTGCGTCTTGGGCAACA 3720
Db 8660 AGACGATGAGAAATTAAGATTCAGACATCTGACCCCTGCGTCTTGGGCAACA 8719
QY 3721 AGGTAATCTCTGTTAGCTCAATGCTCTCAGGTGACAGTGTAACTGATACAGAG 3780
Db 8720 AGGTAATCTCTGTTAGCTCAATGCTCTCAGGTGACAGTGTAACTGATACAGAG 8779
QY 3781 CGGACATCTGAGAAATTCATGACCGGTGAGAAAAAGATCAGAGAAAGTTTGTGGTGA 3840
Db 8780 CGGACATCTGAGAAATTCATGACCGGTGAGAAAAAGATCAGAGAAAGTTTGTGGTGA 8839
QY 3841 AGAGAGATCTTGTCCACCCGCTCAGTGAAGAGTGTAAAGTCCAGTAAAGATCA 3900
Db 8840 AGAGAGATCTTGTCCACCCGCTCAGTGAAGAGTGTAAAGTCCAGTAAAGATCA 8899
QY 3901 CTTGAAGAGACGCTGCGGGTACATACATGACAGGCGGACCAACGCTATTA 3960
Db 8900 CTTGAAGAGACGCTGCGGGTACATACATGACAGGCGGACCAACGCTATTA 8959
QY 3961 GTCTATCTGAGAGAGGCTGAGGGAAGTGTACATTAACCACTTGTGCGAAGACGT 4020
Db 8960 GTCTATCTGAGAGAGGCTGAGGGAAGTGTACATTAACCACTTGTGCGAAGACGT 9019
QY 4021 CACCTAGCAATGTAGTGTGCGACTACAGACAGGTATCTGAGACAGCGGACGAAGAT 4080
Db 9020 CACCTAGCAATGTAGTGTGCGACTACAGACAGGTATCTGAGACAGCGGACGAAGAT 9079
QY 4081 GAAACGGCTGCACTAAAGCAAAACAGTGCATTGCTTCAAGAGCGAACAAAGAAATGGT 4140
Db 9080 GAAACGGCTGCACTAAAGCAAAACAGTGCATTGCTTCAAGAGCGAACAAAGAAATGGT 9139
QY 4141 CTTCAACTGCGCGATCTTATTAAGCACAGACCACTCAGTGCAGGTTAAATGTCACAT 4200
Db 9140 CTTCAACTGCGCGATCTTATTAAGCACAGACCACTCAGTGCAGGTTAAATGTCACAT 9199
QY 4201 TCCATTCGCTTGAACACGACAGTGTGCGGCTTCCGTTAAGTCAAGGCTTACAGTCA 4260
Db 9200 TCCATTCGCTTGAACACGACAGTGTGCGGCTTCCGTTAAGTCAAGGCTTACAGTCA 9259
QY 4261 GAAAGTGTTCAAAGGATCACTCTGCACTGATGCAATGCGAACCAATGCTGACAAC 4320
Db 9260 GAAAGTGTTCAAAGGATCACTCTGCACTGATGCAATGCGAACCAATGCTGACAAC 9319
QY 4321 GAGAAATTTGGGCTGCGAGAGACGAAACAGCAAGATGATTTACAGGCTTACATCAG 4380
Db 9320 GAGAAATTTGGGCTGCGAGAGACGAAACAGCAAGATGATTTACAGGCTTACATCAG 9379
QY 4381 GAAATTTTCTGTGGGCGAGAAAGGCTGAGTACGTATGGGGTAACTTAACCAAGTCA 4440
Db 9380 GAAATTTTCTGTGGGCGAGAAAGGCTGAGTACGTATGGGGTAACTTAACCAAGTCA 9439
QY 4441 AGTCTGGGCCAGAGAGTGGGACACGAGGACCAACATGATGCGCATGATCATCAT 4500

Db	9440	AGTGTGGGCCAGGAGGTGGGCA	CCAGGGCGAC	CCCAATGATGGCCGCATGAGATCATCAT	9499
Qy	4501	CCACTATTTATTCATGCGATCCAGTCTA	CACTGTCAATTTGTGTGTGTGTGTGTGTGC	4560	
Db	9500	CCAATATTTATTCATGCGATCCAGTCTA	CTGTCAATTTGTGTGTGTGTGTGTGTGC	9559	
Qy	4561	TATCTCGGTAGGCACTGCAATCAGAGAGCTTGCATTCGCAAGCAAGAAAGACATGCTCT	4620		
Db	9560	TATCTCGGTAGGCACTGCAATCAGAGAGCTTGCATTCGCAAGCAAGAAAGACATGCTCT	9619		
Qy	4621	GACGCAATACGCGCTTGCACCCGAACGGAACGGTACCACAGCATTAAGCGTATTTGTGTGCTG	4680		
Db	9620	GACGCAATACGCGCTTGCACCCGAACGGAACGGTACCACAGCATTAAGCGTATTTGTGTGCTG	9679		
Qy	4681	CATTGGGCGAACCAACGCTGAAACATTGGAGAAACCTTGAACCATCTGTGTTTAAACA	4740		
Db	9680	CATTGGGCGAACCAACGCTGAAACATTGGAGAAACCTTGAACCATCTGTGTTTAAACA	9739		
Qy	4741	CCAAACGTTTTCTCTGGGCGACAGTTTGTGATTCCTCTGTGACGCGCTTGTTATTTCTGTCCG	4800		
Db	9740	CCAAACGTTTTCTCTGGGCGACAGTTTGTGATTCCTCTGTGACGCGCTTGTTATTTCTGTCCG	9799		
Qy	4801	CTGTCTTTTCATGTGCGATGCGCTTTTATTAATGTTGACAGGCGTGTGCTGTGGGGAAGGTAGA	4860		
Db	9800	CTGTCTTTTCATGTGCGATGCGCTTTTATTAATGTTGACAGGCGTGTGCTGTGGGGAAGGTAGA	9859		
Qy	4861	CGCCTTTGAAATATGCGACACATGTGCCAAATGTTCCGGGGATCCCGTATTAAGCGCTTGTGT	4920		
Db	9860	CGCCTTTGAAATATGCGACACATGTGCCAAATGTTCCGGGGATCCCGTATTAAGCGCTTGTGT	9919		
Qy	4921	CGAAACGGCGAGTTTACGCGCCCATCTTAACCTGTGAGATCAAGGTGTCTCATCGGAATTAAC	4980		
Db	9920	CGAAACGGCGAGTTTACGCGCCCATCTTAACCTGTGAGATCAAGGTGTCTCATCGGAATTAAC	9979		
Qy	4981	ACCTTCAACTTAACCAAGAGTACGTAACCTGCAAAATTCACACAGATCAATTCCTTACACACA	5040		
Db	9980	ACCTTCAACTTAACCAAGAGTACGTAACCTGCAAAATTCACACAGATCAATTCCTTACACACA	10039		
Qy	5041	AGTTAAATGCTGGGGGTCCCTCGAGTGCAGAGGATCCTCAAAAGCGGATTAACAATGCCG	5100		
Db	10040	AGTTAAATGCTGGGGGTCCCTCGAGTGCAGAGGATCCTCAAAAGCGGATTAACAATGCCG	10099		
Qy	5101	CGTTTTTGGGCGGTGTGTACCTCTTTCATGTGGGGAAGGCGCAATGCTTCTGTGACAGTGA	5160		
Db	10100	CGTTTTTGGGCGGTGTGTACCTCTTTCATGTGGGGAAGGCGCAATGCTTCTGTGACAGTGA	10159		
Qy	5161	GAAACACACATGATGTAGGCGGATACGTCGAGTTCCGCTCCAGATGTGACATATGATATGATGCC	5220		
Db	10160	GAAACACACATGATGTAGGCGGATACGTCGAGTTCCGCTCCAGATGTGACATATGATATGATGCC	10219		
Qy	5221	AGTGCACATAAAAGTTTACACAGCTGCTCGAAAGTGCAGCTGCGCTGATATGATATACGGCAA	5280		
Db	10220	AGTGCACATAAAAGTTTACACAGCTGCTCGAAAGTGCAGCTGCGCTGATATGATATACGGCAA	10279		
Qy	5281	CACCAACCGCGACCTGTGATACGTTTGTCAATGTGCGTGCACGCGAGTTCTCTCACGGGACCT	5340		
Db	10280	CACCAACCGCGACCTGTGATACGTTTGTCAATGTGCGTGCACGCGAGTTCTCTCACGGGACCT	10339		
Qy	5341	GAAAGTATATGACAGGGCGGATATCAGCGGCTTTTTCACCTTTGACATATAGGTGTGCTAT	5400		
Db	10340	GAAAGTATATGACAGGGCGGATATCAGCGGCTTTTTCACCTTTGACATATAGGTGTGCTAT	10399		
Qy	5401	CAGAAAGGGGCTGTTTACAACTACGACTTCCCTGTAGATGAGATGATAAAACAGAGAC	5460		
Db	10400	CAGAAAGGGGCTGTTTACAACTACGACTTCCCTGTAGATGAGATGATAAAACAGAGAC	10459		
Qy	5461	GTTTCGGGATATTAAGACATCTTCGCTTGATGTCTACAGACATATGATAGCCCGACCTGACAT	5520		
Db	10460	GTTTCGGGATATTAAGACATCTTCGCTTGATGTCTACAGACATATGATAGCCCGACCTGACAT	10519		
Qy	5521	ACGGCTGTGAGGCTTCTGTCAAGAAATCAGCTGCCCTTACACCAAGCAATATCAAG	5580		

Db	10520	ACGGCTCTGAACCTCTTCTGTCAAGAACATCCACGCGCCCTCAACCCCAAGCAGTATCAGG	10579
Oy	5581	GTATGAAATGTGGAAGAACAACTCAGAGACAGCCCTCTGCAAGAAACAGACCACTTTGATG	5640
Db	10580	GTATGAAATGTGGAAGAACAACTCAGAGACAGCCCTCTGCAAGAAACAGACCACTTTGATG	10639
Oy	5641	TAAAAATTGAATGGAGCTCTCTGGACGCTCTAACTGTGCTTACCTGGGACATCCCTATCTC	5700
Db	10640	TAAAAATTGAATGGAGCTCTCTGGAGCTCTAACTGTGCTTACCTGGGACATCCCTATCTC	10699
Oy	5701	GATTGACATCCCTGATGACAGCTTTTGTGAGATCATCAGAATCACAACAAATTTAGAAAT	5760
Db	10700	GATTGACATCCCTGATGACAGCTTTTGTGAGATCATCAGAATCACAACAAATTTAGAAAT	10759
Oy	5761	TAGCTGACAGTGAAGACACTGCACTTTATCTGCAACTTTGGTGGTCTCTAACATTTACA	5820
Db	10760	TAGCTGACAGTGAAGACACTGCACTTTATCTGCAACTTTGGTGGTCTCTAACATTTACA	10819
Oy	5821	GTACAAAGCTGACAGAGGAGGACATTGTCCAGTTCACCTCCACTCCACAGACAGCTGTTT	5880
Db	10820	GTACAAAGCTGACAGAGGAGGACATTGTCCAATTCACTCCCACTCCACAGACAGCTGTTT	10879
Oy	5881	GAAGGAAGCGACACACATGTGACTGTCCGTGAGGACAGCATTAACATTAATTTAGACATC	5940
Db	10880	GAAGGAAGCGACACACATGTGACTGTCCGTGAGGACAGCATTAACATTAATTTAGACATC	10939
Oy	5941	GAAGCCCAAGCAAAATTTTATAGTTTGGCTATGCGGCAAGAGTCACTCCGAAAGCTGA	6000
Db	10940	GAAGCCCAAGCAAAATTTTATAGTTTGGCTATGCGGCAAGAGTCACTCCGAAAGCTGA	10999
Oy	6001	ATGTAAACCAACCGGCGCGACCAACATAATTTGGAGAACACACATAAGTGCACCAAGATTTCA	6060
Db	11000	ATGTAAACCAACCGGCGCGACCAACATAATTTGGAGAACACACATAAGTGCACCAAGATTTCA	11059
Oy	6061	GGCGGCAGTTTCCAAAACATCTTGGAACTGGCTGCTTGACACTGTTTGGGGAGCATCATC	6120
Db	11060	GGCGGCAGTTTCCAAAACATCTTGGAACTGGCTGCTTGACACTGTTTGGGGAGCATCATC	11119
Oy	6121	CCCATATGTTGTGAGACTTAATAGTGTGGTCTGACAGCTATAGCTTATTAACACAGTAG	6180
Db	11120	CCCATATGTTGTGAGACTTAATAGTGTGGTCTGACAGCTTATAGCTTATTAACACAGTAG	11179
Oy	6181	ATGACTGAGGCGGAGACATGACATAGCGGTAAAACTCGATGTACTTCCGAGGAAACGTGG	6240
Db	11180	ATGACTGAGGCGGAGACATGACATAGCGGTAAAACTCGATGTACTTCCGAGGAAACGTGG	11239
Oy	6241	TGCATAATGTCACAGCGCCGCTTGACACTAAAACTCGATGTATTTCCGAGGAAGCAAGTG	6300
Db	11240	TGCATAATGTCACAGCGCCGCTTGACACTAAAACTCGATGTATTTCCGAGGAAGCAAGTG	11299
Oy	6301	CATTAAGCTGTGAGAGTGCACATTAATGCTAATATACACTAATATTTAACACACTATAT	6360
Db	11300	CATTAAGCTGTGAGAGTGCACATTAATGCTAATATACACTAATATTTAACACACTATAT	11359
Oy	6361	CACCTTTATAGACTACTATAGGCTCTAATATATACATACACATATTTTACTTAAAAAC	6420
Db	11360	CACCTTTATAGACTACTATAGGCTCTAATATATACATACACATATTTTACTTAAAAAC	11419
Oy	6421	ACTAATACACTTTAATTAATCTTTTATAAATTTTCTTTTGTTTTATTTTGTTTTAAA	6480
Db	11420	ACTAATACACTTTAATTAATCTTTTATAAATTTTCTTTTGTTTTATTTTGTTTTAAA	11479
Oy	6481	ATTTC 6485	
Db	11480	ATTTC 11484	

RESULT 2
US-10-023-649-7
; Sequence 7, Application US/10023649
; Publication No. US2003014320A1
; GENERAL INFORMATION:
; APPLICANT: The Minister of National Defence, Government of Canada

```

1 APPLICANT: Nagata, Leslie P
2 APPLICANT: Wong, Jonathan P
3 TITLE OF INVENTION: NO. US20030143201a1el DNA-Based Vaccine Against the Encephalitis
4 FILE REFERENCE: NEI-001
5 CURRENT APPLICATION NUMBER: US/10/023,649
6 CURRENT FILING DATE: 2002-12-27
7 PRIOR APPLICATION NUMBER: 60/256,948
8 PRIOR FILING DATE: 2000-12-21
9 NUMBER OF SEQ ID NOS: 7
10 SOFTWARE: PatentIn version 3.1
11 SEQ ID NO 7
12 LENGTH: 4395
13 TYPE: DNA
14 ORGANISM: western equine encephalomyelitis virus - STRAIN 71v-1658
15 FEATURE:
16 NAME/KEY: CMV promoter
17 LOCATION: (1)..(1260)
18 OTHER INFORMATION: Pax vector sequence: 1-196; CMV promoter: 1-115; CMV putative tr
19 OTHER INFORMATION: anecliptonal seat site: 125; T7 promoter: 48-167; PVAX multicit
20 OTHER INFORMATION: ning region: 168-196; polyprotein (C-E3-E2-6K-E1): 214-4065; pCDM
21 OTHER INFORMATION: -HX45 nontranslated region: 4066-4348; pCDM-HX45 vector sequence
22 OTHER INFORMATION: 4349-4385; PVAX vector sequence: 4386
23
24 -IS-10-023-649-7

```

Query Match	63.8%	Score 4137.2;	DB 12;	Length 4395;
Best Local Similarity	99.9%	Pred. No. 0;		
Matches 4139;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;

QY	2324	CGAGTCACTACTGGCAGGCGCTGATCATCACTCTCTGTCCACGTTAGCCGAAGCGTTAA	2383
Db	207	CGAGCTCACTACTGGCAGGCGCTGATCATCACTCTCTGTCCACGTTAGCCGAAGCGTTAA	266
QY	2384	GAACCTTCAGAGCATTAAGAGGAGGCCAATCAACCTCTACAGCGTCACTTAATAGTAC	2443
Db	267	GAACCTTCAGAGCATTAAGAGGAGGCCAATCAACCTCTACAGCGTCACTTAATAGTAC	326
QY	2444	GTAAGTAGACACGCACCTTACCCACCGGCAGAAATGTTTCCATACCTCACTGAACCTTTCCA	2503
Db	327	GTAAGTAGACACGCACCTTACCCACCGGCAGAAATGTTTCCATACCTCACTGAACCTTTCCA	386
QY	2504	CCAGTTTACCTCAACAAATCCGATGGCTTACCGAGATTCGAAACCCCTCTAGGCGCGCTGG	2563
Db	387	CCAGTTTACCTCAACAAATCCGATGGCTTACCGAGATTCGAAACCCCTCTAGGCGCGCTGG	446
QY	2564	AGGCGGTTTGGGCCCCCGCTGGCTGCTCAAAATCGAAGATCTTAGAGGTCGATAGTCAAC	2623
Db	447	AGGCGGTTTGGGCCCCCGCTGGCTGCTCAAAATCGAAGATCTTAGAGGTCGATAGTCAAC	506
QY	2624	TTGACTTTCCAAACAACGATTCACCTTAATCCGCGCGCAAGTCCACCGGCCAAAGAAAGAAAG	2683
Db	507	TTGACTTTCCAAACAACGATTCACCTTAATCCGCGCGCAAGTCCACCGGCCAAAGAAAGAAAG	566
QY	2684	AGTGTCTCTTAAGCAAAAACCTACTCAGCCTTAATAAAGAAAGACGACCAAGAGGACG	2743
Db	567	AGTGTCTCTTAAGCAAAAACCTACTCAGCCTTAATAAAGAAAGACGACCAAGAGGACG	626
QY	2744	AAAGCGCAAGCTTAACAACGAGGAAACGACACGATGTGTATGAAAGTTGAGTCCGACAAAG	2803
Db	627	AAAGCGCAAGCTTAACAACGAGGAAACGACACGATGTGTATGAAAGTTGAGTCCGACAAAG	686
QY	2804	ACATTTTCGATCATGCTGTAAGCGCCAAAGTGAATGGAATGTGCTGCGTGTGCGAGGAAG	2863
Db	687	ACATTTTCGATCATGCTGTAAGCGCCAAAGTGAATGGAATGTGCTGCGTGTGCGAGGAAG	746
QY	2864	CTGATGAACCACTTCACGCTTGAAAGAAAAATTGATTAATGACAAATTAGCGGCGGTGAA	2923
Db	747	CTGATGAACCACTTCACGCTTGAAAGAAAAATTGATTAATGACAAATTAGCGGCGGTGAA	806
QY	2924	TTGAAGAAGGCTAGCATGTACGACTTGAGAGTACGGCGAGCTTTCCACAGAAATGAAATCA	2983
Db	807	TTGAAGAAGGCTAGCATGTACGACTTGAGAGTACGGCGAGCTTTCCACAGAAATGAAATCA	866
QY	2984	GACACGCTGCAGTACACCAAGCGACAAACCAACCGGCGTTCTTCAACATGGGACCAACGCGCGCA	3043

Db	867	GACACGCTGCAGTACACACGGACAACACACCGGGCTTCTCAACTGGCACCACACGGCCCA	926
QY	3044	GTCCAGTATGAAATGGAGATTATACCTTACCGAGAGAGTGGCGGAAAGCCAGACAC	3103
Db	927	GTCCAGTATGAAATGGAGATTATACCTTACCGAGAGAGTGGCGGAAAGCCAGACAC	986
QY	3104	GGAAACCGGATCTTGAGCAACAGAGCGAGTGTGGCTATGTGTTAGAGGTGCAAT	3163
Db	987	GGAAACCGGATCTTGAGCAACAGAGCGAGTGTGGCTATGTGTTAGAGGTGCAAT	1046
QY	3164	GAGGGCAGCGCTACGGCGCTTTCAGTGTCACTTGAAACAGAAAGGGGTGACCTTATGG	3222
Db	1047	GAGGGCAGCGCTACGGCGCTTTCAGTGTCACTTGAAACAGAAAGGGGTGACCTTATGG	1106
QY	3224	GATACCCCCGAAAGGTTCTGAACCGGTGTCATAGTACAGGCGTATGGGTGTTTCGAT	3283
Db	1107	GATACCCCCGAAAGGTTCTGAACCGGTGTCATAGTACAGGCGTATGGGTGTTTCGAT	1166
QY	3284	GTCACGTTCCATGCGACAAACCAACCCGTGTGCTATTCATCTGACGCCAGAAAGCACTC	3343
Db	1167	GTCACGTTCCATGCGACAAACCAACCCGTGTGCTATTCATCTGACGCCAGAAAGCACTC	1222
QY	3344	GACGTGCTCGAAGAGAACTGACATTCAAATTCAGACAGCTGCTGGAGAACGTCTTG	3403
Db	1227	GACGTGCTCGAAGAGAACTGACATTCAAATTCAGACAGCTGCTGGAGAACGTCTTG	1286
QY	3404	AAATGTCATACGCGCGGCCCAAGCAAGCAATTCGCGTTCCTCACCTGACCGAGCTCC	3463
Db	1287	AAATGTCATACGCGCGGCCCAAGCAATTCGCGTTCCTCACCTGACCGAGCTCC	1344
QY	3464	TACCTGGGGTTCGCGCGTATTTGGACACATCCACGCCGTGTTTCAGCCCAATAAAT	3522
Db	1347	TACCTGGGGTTCGCGCGTATTTGGACACATCCACGCCGTGTTTCAGCCCAATAAAT	1406
QY	3524	GAGAACGTGTGGAGCAATCTGATGATGATTCGATTAGATTCAGGTTCCGACACATTC	3583
Db	1407	GAGAACGTGTGGAGCAATCTGATGATGATTCGATTAGATTCAGGTTCCGACACATTC	1466
QY	3584	GGCTACATCAGGAGGACCTGCGGATGTCAACAAATTCGCTTCAATGCTTTTGGACAC	3643
Db	1467	GGCTACATCAGGAGGACCTGCGGATGTCAACAAATTCGCTTCAATGCTTTTGGACAC	1522
QY	3644	GACCATGACATCAAGAAAGACATGATGAGAAATAGCTATCAGACACATCTGCACCTGC	3703
Db	1527	GACCATGACATCAAGAAAGACATGATGAGAAATAGCTATCAGACACATCTGCACCTGC	1586
QY	3704	CGTCGCTTGGCCACAAAGGGTACTTCTGTTAGCTCAATGTCTCCAGGTGACAGTGA	3763
Db	1587	CGTCGCTTGGCCACAAAGGGTACTTCTGTTAGCTCAATGTCTCCAGGTGACAGTGA	1646
QY	3764	ACCGTCAGTATCAAGAGCGGAGCATCTGGAATTCATGCAACCGTGGAGAAAAAATCATCG	3822
Db	1647	ACCGTCAGTATCAAGAGCGGAGCATCTGGAATTCATGCAACCGTGGAGAAAAAATCATCG	1706
QY	3824	AGGAAGTTGTCCGCTAGAGAGAGTACTTGTCCACCCGTCACATGAAAGCTGGTAAAG	3883
Db	1707	AGGAAGTTGTCCGCTAGAGAGAGTACTTGTCCACCCGTCACATGAAAGCTGGTAAAG	1766
QY	3884	TGCCACGTTTACGATCACTTGAAAGAGACGTCTCCGGGTACATTAACCATGACACAGGCCA	3943
Db	1767	TGCCACGTTTACGATCACTTGAAAGAGACGTCTCCGGGTACATTAACCATGACACAGGCCA	1822
QY	3944	GGCCCAACCGGTATTAAGTCTTACTTGAGGAAACGCTCAGGCGGAAAGTGAATTAACA	4003
Db	1827	GGCCCAACCGGTATTAAGTCTTACTTGAGGAAACGCTCAGGCGGAAAGTGAATTAACA	1886
QY	4004	CCCTTGGCAAGAACGTCACTTACGAATGTAAAGTGGCGCATCAGACACAGTATCGTG	4063
Db	1887	CCCTTGGCAAGAACGTCACTTACGAATGTAAAGTGGCGCATCAGACACAGTATCGTG	1946
QY	4064	AGCAGCGGAACGAAGATGAACGCTGCACTTAAGCAAAACGTGCATTTGCTTACAAAGCC	4123

Dh 1947 AGCAGCGAAGCAAGATGAAAGCGCTGCATTAAGCAAAAGATGCTATGGCTCAAGAGC 2006
Qy 4124 GACCAAGCAAAATGGGCTTCACTCGCCGGATGCTTATGAGCAACAGACCACTGAGT 4183
Dh 2007 GACCAAGCAAAATGGGCTTCACTCGCCGGATGCTTATGAGCAACAGACCACTGAGT 2066
Qy 4184 CAAGGTAAATGTCATTCATTCGCTTGACACCGACAGTCTGCCGCTTCCGTTAGCT 4243
Dh 2067 CAAGGTAAATGTCATTCATTCGCTTGACACCGACAGTCTGCCGCTTCCGTTAGCT 2126
Qy 4244 CACACGCTTCAAGTCAAGAAAGTGTCAAAAGGATCAACCTTCCACTGATGCAATCGA 4303
Dh 2127 CACACGCTTCAAGTCAAGAAAGTGTCAAAAGGATCAACCTTCCACTGATGCAATCGA 2186
Qy 4304 CCAACATGCTGCAAGAGAAATGGGGCTGAGACAGACGCAAGAGAAATGGAAT 4363
Dh 2187 CCAACATGCTGCAAGAGAAATGGGGCTGAGACAGACGCAAGAGAAATGGAAT 2246
Qy 4364 ACAGGCTCTACATCAAGAAATTTTCTGTGGGCGAAGAGGCTGAGATAGTATGGGCT 4423
Dh 2247 ACAGGCTCTACATCAAGAAATTTTCTGTGGGCGAAGAGGCTGAGATAGTATGGGCT 2306
Qy 4424 AACCATGAACAGTCAAGTCTGGGCGCAGAGATGCGACACCGACCAATGATGG 4483
Dh 2307 AACCATGAACAGTCAAGTCTGGGCGCAGAGATGCGACACCGACCAATGATGG 2366
Qy 4484 CCGCATGATGATCAATCCATATATCATGGGATCAGTCTACATGCTATGCTGCTG 4543
Dh 2367 CCGCATGATGATCAATCCATATATCATGGGATCAGTCTACATGCTATGCTGCTG 2426
Qy 4544 TGTGCTGCTCTTGTCTATCTGTGTAGAGCACTGATCATCAAGAGCTTGATGCGCAA 4603
Dh 2427 TGTGCTGCTCTTGTCTATCTGTGTAGAGCACTGATCATCAAGAGCTTGATGCGCAA 2486
Qy 4604 GCAAGAAAGATGCTGCTGAGCGCCATACGCGCTTGCAACCGAACGCAACGCAACGA 4663
Dh 2487 GCAAGAAAGATGCTGCTGAGCGCCATACGCGCTTGCAACCGAACGCAACGCAACGA 2546
Qy 4664 TTAGCGGTTTTGTGCTGATTCGCGCAACCAAGCTGAAACATTTGAGAACTTTGAAAC 4723
Dh 2547 TTAGCGGTTTTGTGCTGATTCGCGCAACCAAGCTGAAACATTTGAGAACTTTGAAAC 2606
Qy 4724 CATCTGTGTTTAAACAACAACGCTTCTCTGCGCAGAGTGTGATCTCTCTGCGACGC 4783
Dh 2607 CATCTGTGTTTAAACAACAACGCTTCTCTGCGCAGAGTGTGATCTCTCTGCGACGC 2666
Qy 4784 CTGTGTAATCTGCTTCCGCTGCTTTCATGCTGATGCTCTTTTATTTGTTGACGCGTC 4843
Dh 2667 CTGTGTAATCTGCTTCCGCTGCTTTCATGCTGATGCTCTTTTATTTGTTGACGCGTC 2726
Qy 4844 TGCGTGGGGAAGTGAAGCGCTTGGAACATGCGACCACTGCGCAAAATGTTCCGGGATC 4903
Dh 2727 TGCGTGGGGAAGTGAAGCGCTTGGAACATGCGACCACTGCGCAAAATGTTCCGGGATC 2786
Qy 4904 CCGATTAAGGCTGTGTCGACGCGAGTTACGCGCACTTAACTGAGATCAAGCTC 4963
Dh 2787 CCGATTAAGGCTGTGTCGACGCGAGTTACGCGCACTTAACTGAGATCAAGCTC 2846
Qy 4964 GTCTCATCGGAATTAACAACCTTCAACTTAACAAGAGATGCTGACCTGCAAAATTCACACA 5023
Dh 2847 GTCTCATCGGAATTAACAACCTTCAACTTAACAAGAGATGCTGACCTGCAAAATTCACACA 2906
Qy 5024 GTCAATTCCTTCAACAAGTAATGCTGCGGGTCCCTCGAGTGCAGAGGATCTCTCAAG 5083
Dh 2907 GTCAATTCCTTCAACAAGTAATGCTGCGGGTCCCTCGAGTGCAGAGGATCTCTCAAG 2966
Qy 5084 GCGGATTAACATGCGCGGTTTTTGGCGGTGTGTAACCTTTCATGTGGGAGCGCAAA 5143
Dh 2967 GCGGATTAACATGCGCGGTTTTTGGCGGTGTGTAACCTTTCATGTGGGAGCGCAAA 3026
Qy 5144 TGCTTCTGTGACAGTGAACAACAACACTGATGAGGCGTATGCTGAGTTGCTTCAGAC 5203
Dh 3027 TGCTTCTGTGACAGTGAACAACAACACTGATGAGGCGTATGCTGAGTTGCTTCAGAC 3086

Qy 5204 TGCATATAGATCAGCAGTCCGACTTAAAGTTTCAACACAGCTGCTGAAAGTCCGCTG 5263
Dh 3087 TGCATATAGATCAGCAGTCCGACTTAAAGTTTCAACACAGCTGCTGAAAGTCCGCTG 3146
Qy 5264 CGTATAGTATACGGCAACACACCGCGCATCTGGAATCGTTTGTCAAATGGCGTCAAGCA 5323
Dh 3147 CGTATAGTATACGGCAACACACCGCGCATCTGGAATCGTTTGTCAAATGGCGTCAAGCA 3206
Qy 5324 GGTTCCTCAGGGACCTGAAGGTCATAGCAGGGCCGATACACCGCTTTTACCCCTT 5383
Dh 3207 GGTTCCTCAGGGACCTGAAGGTCATAGCAGGGCCGATACACCGCTTTTACCCCTT 3266
Qy 5384 GACCATTAAGTCTGATCAGAAAGGGCTGTTTAACTACATCACTTCCCTGATATGA 5443
Dh 3267 GACCATTAAGTCTGATCAGAAAGGGCTGTTTAACTACATCACTTCCCTGATATGA 3326
Qy 5444 GCTATGAACACAGAGGCTTCCGATATTAAGCATCTCTGCTTGAATGCTACAGACATA 5503
Dh 3327 GCTATGAACACAGAGGCTTCCGATATTAAGCATCTCTGCTTGAATGCTACAGACATA 3386
Qy 5504 GTAGCCCGCATGACATACGCGCTGCTGAAGCCTTCTGTCAAGAACATCCACGTCCTAC 5563
Dh 3387 GTAGCCCGCATGACATACGCGCTGCTGAAGCCTTCTGTCAAGAACATCCACGTCCTAC 3446
Qy 5564 ACCCAAGCAGTATCAGGATATGAATGTGAAAGAACAACTCAGAGACACCCCTGCAAGAA 5623
Dh 3447 ACCCAAGCAGTATCAGGATATGAATGTGAAAGAACAACTCAGAGACACCCCTGCAAGAA 3506
Qy 5624 ACAGACCAATTTGATGATTAATTAATGAAGTGAAGCTCTGCGACGCTTAACTGCTTAC 5683
Dh 3507 ACAGACCAATTTGATGATTAATTAATGAAGTGAAGCTCTGCGACGCTTAACTGCTTAC 3566
Qy 5684 GGGCAATCCCTATCTCGATTTGACATCCCTGATGAGCCTTTGAGATATATAGAAATCA 5743
Dh 3567 GGGCAATCCCTATCTCGATTTGACATCCCTGATGAGCCTTTGAGATATATAGAAATCA 3626
Qy 5744 CCAACATTTTGAAGATTTAGCTGACAGTACAGACTGCACTTTATTTGCGACATTTGGT 5803
Dh 3627 CCAACATTTTGAAGATTTAGCTGACAGTACAGACTGCACTTTATTTGCGACATTTGGT 3686
Qy 5804 GGTTCCTTAACTTAACATTAAGTGAAGAGCTGACAGGAGGACATTTGCTACTCCAC 5863
Dh 3687 GGTTCCTTAACTTAACATTAAGTGAAGAGCTGACAGGAGGACATTTGCTACTCCAC 3746
Qy 5864 TCCACGACAGCTGTTTGAAGAGGACCAACATGTCACGCTGAGGACAGATACA 5923
Dh 3747 TCCACGACAGCTGTTTGAAGAGGACCAACATGTCACGCTGAGGACAGATACA 3806
Qy 5924 CTACATTTTGAACATGACCCCAACAAGCAAAATTTTATGTTTGCATGCGCAAGAG 5983
Dh 3807 CTACATTTTGAACATGACCCCAACAAGCAAAATTTTATGTTTGCATGCGCAAGAG 3866
Qy 5984 TCCACTGCAATGCTGAATGTAACCAACGCGCGCACCAATTAATGAGAACCAATPAA 6043
Dh 3867 TCCACTGCAATGCTGAATGTAACCAACGCGCGCACCAATTAATGAGAACCAATPAA 3926
Qy 6044 GTGACCAAAATTTCCAGGCGGCAAGTTTCCAAAACATCTTGAACCTGCTGCACTG 6103
Dh 3927 GTGACCAAAATTTCCAGGCGGCAAGTTTCCAAAACATCTTGAACCTGCTGCACTG 3986
Qy 6104 TTTGGGGGACATCATCTCTCATTTGTTGATGAGCTTATGATGTTGCTGCAAGCTATG 6163
Dh 3987 TTTGGGGGACATCATCTCTCATTTGTTGATGAGCTTATGATGTTGCTGCAAGCTATG 4046
Qy 6164 CTATATAACACAGTATAGATCTGAGCGCGAACAATGACATAGCGGTAATCTGATGTA 6223
Dh 4047 CTATATAACACAGTATAGATCTGAGCGCGAACAATGACATAGCGGTAATCTGATGTA 4106
Qy 6224 CTTCCGAGGAAGCTGTGATATAATGCCAGCGCCGCTTGAACATGCAATGATATT 6283
Dh 4107 CTTCCGAGGAAGCTGTGATATAATGCCAGCGCCGCTTGAACATGCAATGATATT 4166

6284 TCCGAGGAGCAGCTGATTAATGCTGTGACAGTGCACATTAATGATATACACTACA 6343
4167 TCCGAGGAGCAGCTGATTAATGCTGTGACAGTGCACATTAATGATATACACTACA 4226
6344 TATTAAACAACCTATATCACTTTTATGAGACTCACTAATGCTCTTAATATACACTACAC 6403
4227 TATTAAACAACCTATATCACTTTTATGAGACTCACTAATGCTCTTAATATACACTACAC 4286
6404 ATATTTCCTTAAAAAAGCTATACACCTTATTAATCTTTATTAATTTTCTTTTGT 6463
4287 ATATTTCCTTAAAAAAGCTATACACCTTATTAATCTTTTATTAATTTTCTTTTGT 4346
6464 TT 6465
4347 TT 4348

RESULT 3
US-10-023-649-5
Sequence 5, Application US/10023649
Publication No. US20030143201A1
GENERAL INFORMATION:
APPLICANT: The Minister of National Defence, Government of Canada
APPLICANT: Negata, Leslie P
TITLE OF INVENTION: No. US20030143201A1el DNA-Based Vaccine Against the Encephalitis
FILE REFERENCE: NEL-001
CURRENT APPLICATION NUMBER: US/10/023, 649
CURRENT FILING DATE: 2002-12-27
PRIOR APPLICATION NUMBER: 60/256,948
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 4150
TYPE: DNA
ORGANISM: Western equine encephalomyelitis virus - strain 71V-1658
FEATURE:
NAME/KEY: CDS
LOCATION: (159)..(3869)
OTHER INFORMATION: vector sequence 1-9; 5' SacI primer 9-20; 3' end of NS4 gene 16-1
OTHER INFORMATION: 14; intergenic region 115-158; polypeptide (C-E3-E2-6K-E1) 159-38
OTHER INFORMATION: 56; pCDM-XH7 nontranslated region 3857-4150
US-10-023-649-5

Query Match 63.7%; Score 4129.8; DB 12; Length 4150;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 4134; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

2325 GAGATCACTAGGCGGCTGATCATCAAGTCTGTCCAGCTTAGCCGAAAGCGTTAAG 2384
10 GAGCTCACTAGGCGGCTGATCATCAAGTCTGTCCAGCTTAGCCGAAAGCGTTAAG 69
2385 AACTTCAAGAGCATTAAGAGGAGGCCAATCACTCTCTAAGGCTGAAGCTTAATAGGTGAG 2444
70 AACTTCAAGAGCATTAAGAGGAGGCCAATCACTCTCTAAGGCTGAAGCTTAATAGGTGAG 129
2445 TAGTGAACAGCAGCTTACCCACCGGAGAAATTTTTCATACCTTACGTGAAGCTTTCCAC 2504
130 TAGTGAACAGCAGCTTACCCACCGGAGAAATTTTTCATACCTTACGTGAAGCTTTCCAC 189
2505 CAGTTTACCTTACCAAAATCGATGGCTTACCGAGATCCAAACCTCTCTAGGGCGGCTGGA 2564
190 CAGTTTACCTTACCAAAATCGATGGCTTACCGAGATCCAAACCTCTCTAGGGCGGCTGGA 249
2565 GGGCGTTTGGCGCGGCTGCTGCTCAATTCGAAGATCTTATGAGAGTCAATGTAAGTCACT 2624
250 GGGCGTTTGGCGCGGCTGCTGCTCAATTCGAAGATCTTATGAGAGTCAATGTAAGTCACT 309
2625 TGACTTTCAAAAGATCACTTAATTCGGCGGCGGCTGCAAGCTTGAAGAGTCAAGTCAAGT 2684
310 TGACTTTCAAAAGATCACTTAATTCGGCGGCGGCTGCAAGCTTGAAGAGTCAAGTCAAGT 369

2685 GTGCTCTTAAGCAAAACCTTACTAGCTTAAAAAGAGAGAGCAAGCAAGAGAGCA 2744
370 GTGCTCTTAAGCAAAACCTTACTAGCTTAAAAAGAGAGAGCAAGCAAGAGAGCA 429
2745 AACGAGCCTTAAACAGAGAGAGAGCAAGCAAGTATGTATGAAGTTGAGAGTGGACAGCA 2804
430 AACGAGCCTTAAACAGAGAGAGAGCAAGCAAGTATGTATGAAGTTGAGAGTGGACAGCA 489
2805 CATTTCCGATATGATGTAACGGCCAGAGTAAGTAAGTCCGCTGCTTCCGAGAGAGCC 2864
490 CATTTCCGATATGATGTAACGGCCAGAGTAAGTAAGTCCGCTGCTTCCGAGAGAGCC 549
2865 TGATGAACCACTCCACCTTGAAGAGAAATTTGATATAGCAATTTAGCCGCGGTGAAT 2924
550 TGATGAACCACTCCACCTTGAAGAGAAATTTGATATAGCAATTTAGCCGCGGTGAAT 609
2925 TGAAGAGGCTTACGATGTAAGCACTTGAAGTACGGGAGCTTCCCGAGAACATGAATCAG 2984
610 TGAAGAGGCTTACGATGTAAGCACTTGAAGTACGGGAGCTTCCCGAGAACATGAATCAG 669
2985 ACACGCTGAGTACACACGAGAGCAAAACACCGGCTTCTACAACTGGACACCGGCGAG 3044
670 ACACGCTGAGTACACACGAGAGCAAAACACCGGCTTCTACAACTGGACACCGGCGAG 729
3045 TCCAGTATGAGAAATGGAGATTTTACCGTACGAGAGAGTGGCGGAGAAAGCGACAGCG 3104
730 TCCAGTATGAGAAATGGAGATTTTACCGTACGAGAGAGTGGCGGAGAAAGCGACAGCG 789
3105 GAAGACCGATCTTGAGACAAAGAGGAGAGTGTGCTATTTCTTACGAGGTGCAATG 3164
790 GAAGACCGATCTTGAGACAAAGAGGAGAGTGTGCTATTTCTTACGAGGTGCAATG 849
3165 AGGAGACCGCTTACGAGGCTTCTGAGTGTCTTGAACCAAGAAAGGCTGACATTAAGG 3224
850 AGGAGACCGCTTACGAGGCTTCTGAGTGTCTTGAACCAAGAAAGGCTGACATTAAGG 909
3225 ATACCCCGAGAGTCTTGAACCGTGTGCTACTAGTTACAGCGCTATGCTTTCGAATG 3284
910 ATACCCCGAGAGTCTTGAACCGTGTGCTACTAGTTACAGCGCTATGCTTTCGAATG 969
3285 TCAGTTCCTTCCATGCGACAAACACCGCTGTCTATTCCTGACGCGCAGAGCAAGCACTCG 3344
970 TCAGTTCCTTCCATGCGACAAACACCGCTGTCTATTCCTGACGCGCAGAGCAAGCACTCG 1029
3345 ACGTCTGAGAGAGAGTCTGACAACTGACAAATTCAGACAGCTGTGAGAGAGCTTGA 3404
1030 ACGTCTGAGAGAGAGTCTGACAACTGACAAATTCAGACAGCTGTGAGAGAGCTTGA 1089
3405 AATGTCATCAGCGCGGCGCCAAAGAGCAATTTACGATGATCTTCACTGACAGTCCCT 3464
1090 AATGTCATCAGCGCGGCGCCAAAGAGCAATTTACGATGATCTTCACTGACAGTCCCT 1149
3465 AACTGGGTTTGCCTGATTTGACAGACCTCAAGCCGCTTTCAGCCCAATTAATAATG 3524
1150 AACTGGGTTTGCCTGATTTGACAGACCTCAAGCCGCTTTCAGCCCAATTAATAATG 1209
3525 AGAAGTGTGGAGAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3584
1210 AGAAGTGTGGAGAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1269
3585 GCTACAAATCAGAGAGGAGCTGCGGATGTCACCAATTCGTTACATGCTTTTCGACAG 3644
1270 GCTACAAATCAGAGAGGAGCTGCGGATGTCACCAATTCGTTACATGCTTTTCGACAG 1329
3645 ACCATGATCAGAGAGAGAGATGAGAGAAATAGCTATGAGCAATCTGAGACCTGCC 3704
1330 ACCATGATCAGAGAGAGAGATGAGAGAAATAGCTATGAGCAATCTGAGACCTGCC 1389
3705 GTGCTCTTGGGCAAGAGGATCTTCTGTTAGCTCAATGCTTCCAGGTGACAGTGTGA 3764
1390 GTGCTCTTGGGCAAGAGGATCTTCTGTTAGCTCAATGCTTCCAGGTGACAGTGTGA 1449
3765 CCGTCAGTATCAGAGGCGAGCATCTGAGATTCATGACCGTGGAGAAAGATCAGGA 3824

1450 CCGTCAATATACAGAGGGAGCATCTGAAATTCATGACCGTGGAGAAAAGATCAGGA 1509
Qy 3825 GGAAGTTTGTTCGGTAGAGAGAGTACTTTTCCACCCGTCATAGAAAAGTGGTAAAGT 3884
Db 1510 GGAAGTTTGTTCGGTAGAGAGAGTACTTTTCCACCCGTCATAGAAAAGTGGTAAAGT 1569
Qy 3885 GCCACGTTTACATCATTGAAAGGAGAGTCTGCGGGGTACATTAACATGACAGGGCAG 3944
Db 1570 GCCACGTTTACATCATTGAAAGGAGAGTCTGCGGGGTACATTAACATGACAGGGCAG 1629
Qy 3945 GCCCACAACGCGTATAGTCTTATCTGAGAGAGAGCGTCAGGCGAAGTGTACATTAAACAC 4004
Db 1630 GCCCACAACGCGTATAGTCTTATCTGAGAGAGAGCGTCAGGCGAAGTGTACATTAAACAC 1689
Qy 4005 CTTTGGCAAGAAAGTCACTTACGAATGTAAAGTGTGACATACAGCAGAGTATCTGTA 4064
Db 1690 CTTTGGCAAGAAAGTCACTTACGAATGTAAAGTGTGACATACAGCAGAGTATCTGTA 1749
Qy 4065 GCAACGCAAGAAAGTACAGGCGTGCCTAAAGCAAAAGTGTGCTTACCAAGAGCG 4124
Db 1750 GCAACGCAAGAAAGTACAGGCGTGCCTTAAAGCAAAAGTGTGCTTACCAAGAGCG 1809
Qy 4125 ACCAAAGAAATGGGCTTCAACTGCGCGATCTTATTAAGCACAAGACCACTCAAGTGC 4184
Db 1810 ACCAAAGAAATGGGCTTCAACTGCGCGATCTTATTAAGCACAAGACCACTCAAGTGC 1869
Qy 4185 AAGGTAAATTTGACATTCATTCGGCTTGACACCGACAGTCTGCGCGGCTTCTTAACTC 4244
Db 1870 AAGGTAAATTTGACATTCATTCGGCTTGACACCGACAGTCTGCGCGGCTTCTTAACTC 1929
Qy 4245 ACAAGCTTACAGTACAGAGTGGTTCAAGAGCATCACCTTCACTGACCTGCAATGGGAC 4304
Db 1930 ACAAGCTTACAGTACAGAGTGGTTCAAGAGCATCACCTTCACTGACCTGCAATGGGAC 1989
Qy 4305 CAACATTTGCTGACAAACAGAAAATTTGGGGCTGCGAGACAGCAACAGCAAGATGATTA 4364
Db 1990 CAACATTTGCTGACAAACAGAAAATTTGGGGCTGCGAGACAGCAACAGCAAGATGATTA 2049
Qy 4365 CAGGCTTACATCCAGAAATTTTCTGTGGGGCAGAAAGGCTTGAATGATGAGGTA 4424
Db 2050 CAGGCTTACATCCAGAAATTTTCTGTGGGGCAGAAAGGCTTGAATGATGAGGTA 2109
Qy 4425 ACCATGAACAGTACAGAGTGGGCGCAGAGAGTGGGACCGACCGACCCACATGATGGC 4484
Db 2110 ACCATGAACAGTACAGAGTGGGCGCAGAGAGTGGGACCGACCGACCCACATGATGGC 2169
Qy 4485 CGCATGATCATCATCATCTATTTATCATCGGCATCCAGTCTACATGTCTCATTTGCTGT 4544
Db 2170 CGCATGATCATCATCATCTATTTATCATCGGCATCCAGTCTACATGTCTCATTTGCTGT 2229
Qy 4545 GTGTGTCTGCTCTTGTCTATCTGTGAGGACATCATACAGCAGTTTGCATCCCAAG 4604
Db 2230 GTGTGTCTGCTCTTGTCTATCTGTGAGGACATCATACAGCAGTTTGCATCCCAAG 2289
Qy 4605 CAAGAAGACCTGCTACAGCGCATAGCGGCTTGACCGAAGCGAAGGTAACCCACAGCAT 4664
Db 2290 CAAGAAGACCTGCTACAGCGCATAGCGGCTTGACCGAAGCGAAGGTAACCCACAGCAT 2349
Qy 4665 TAGCGTTTGTGTGCTATTCGCGCAACCAACGCTGAAACATTTGGAGAAACCTTGAAC 4724
Db 2350 TAGCGTTTGTGTGCTATTCGCGCAACCAACGCTGAAACATTTGGAGAAACCTTGAAC 2409
Qy 4725 ATCTGTGTTTAAACACCAACCGTTTCTCTGGGACACAGTTGTGATTCCTGCGACGC 4784
Db 2410 ATCTGTGTTTAAACACCAACCGTTTCTCTGGGACACAGTTGTGATTCCTGCGACGC 2469
Qy 4785 TTGTATTCTGTTCCGCTTTCATGCTGACAGCTTTTATTTAGGTTGCAAGGCT 4844
Db 2470 TTGTATTCTGTTCCGCTTTCATGCTGACAGCTTTTATTTAGGTTGCAAGGCT 2529
Qy 4845 GCCTGGGAAGTGAAGCCTTCGAACATGCAACACTGTGCAAAATGTTCCGGGATCC 4904

2530 GCCTGGGAAGTGAAGCCTTCGAACATGCAACACTGTGCAAAATGTTCCGGGATCC 2589
Qy 4905 CGTATAAGGCTGTGTCGAACCGCGAGTTAACCGGCACCTTAACCTGAGATCACGGTGC 4964
Db 2590 CGTATAAGGCTGTGTCGAACCGCGAGTTAACCGGCACCTTAACCTGAGATCACGGTGC 2649
Qy 4965 TCTCATGGAATTAACAACCTTCAACTTAACAAGAGTACGTACCTGCAAAATTCACACAG 5024
Db 2650 TCTCATGGAATTAACAACCTTCAACTTAACAAGAGTACGTACCTGCAAAATTCACACAG 2709
Qy 5025 TCAATCTTCAACAACAAGTTAAATGCTGCGGGTCCCTGAGTGCAAGGATCTCTCAAGG 5084
Db 2710 TCAATCTTCAACAACAAGTTAAATGCTGCGGGTCCCTGAGTGCAAGGATCTCTCAAGG 2769
Qy 5085 CGATTAACAATCCCGGTTTGGCGGTGTGATACCTTTCAATGGGGGAGGCAAT 5144
Db 2770 CGATTAACAATCCCGGTTTGGCGGTGTGATACCTTTCAATGGGGGAGGCAAT 2829
Qy 5145 GCTTCTGTGACATGAGAACACACAACCTGAGAGGCTGAGTGCAGTTCAGACT 5204
Db 2830 GCTTCTGTGACATGAGAACACACAACCTGAGAGGCTGAGTGCAGTTCAGACT 2889
Qy 5205 GCACTATAGATCACGAGTGCCTAAAGTTCAACACAGCTGCTTGAAGTGGCTGC 5264
Db 2890 GCACTATAGATCACGAGTGCCTAAAGTTCAACACAGCTGCTTGAAGTGGCTGC 2949
Qy 5265 GTATAGTATACGGCAACACACCGCGCACTGATACGCTTGTCAATGGGCTACGCGCAG 5324
Db 2950 GTATAGTATACGGCAACACACCGCGCACTGATACGCTTGTCAATGGGCTACGCGCAG 3009
Qy 5325 GTTCTTCAACGGGACCTGAAGGTATAGCAGGCGGATACAGCGCTTTTCAACCTTTG 5384
Db 3010 GTTCTTCAACGGGACCTGAAGGTATAGCAGGCGGATACAGCGCTTTTCAACCTTTG 3069
Qy 5385 ACCATTAAGTGTGATCAGAAAGGCTTTTAACTAGACTTCCTGAGTATGAG 5444
Db 3070 ACCATTAAGTGTGATCAGAAAGGCTTTTAACTAGACTTCCTGAGTATGAG 3129
Qy 5445 CTATGAACCAAGAGGCTTGGGATTTCAAGCATCTGCTGATGCTACAGACATAG 5504
Db 3130 CTATGAACCAAGAGGCTTGGGATTTCAAGCATCTGCTGATGCTACAGACATAG 3189
Qy 5505 TAAGCCGCACTGACATACGCGCTCTGAGGCTTCTGTCAAGAACATTCACAGTCCCTTACA 5564
Db 3190 TAAGCCGCACTGACATACGCGCTCTGAGGCTTCTGTCAAGAACATTCACAGTCCCTTACA 3249
Qy 5565 CCCAAGCATATACAGGATGATAAATGTAAGAAACAACCTAGACACCCCTGCAAGAAA 5624
Db 3250 CCCAAGCATATACAGGATGATAAATGTAAGAAACAACCTAGACACCCCTGCAAGAAA 3309
Qy 5625 CAGACCATTTGATGTAATTAAGTGAAGCTGCGAGGCGTCAACTGTGCTTACG 5684
Db 3310 CAGACCATTTGATGTAATTAAGTGAAGCTGCGAGGCGTCAACTGTGCTTACG 3369
Qy 5685 GGCACATCCCTATCTGATGATGATCCCTGATGAGCTTTTGTGATCATCAGAAATCAC 5744
Db 3370 GGCACATCCCTATCTGATGATGATCCCTGATGAGCTTTTGTGATCATCAGAAATCAC 3429
Qy 5745 CAACAAATTTAGAAATGAGTGCACAGTACAGTGCATTTATTTGCGACAGCTTGGTG 5804
Db 3430 CAACAAATTTAGAAATGAGTGCACAGTACAGTGCATTTATTTGCGACAGCTTGGTG 3489
Qy 5805 GTTCTCTTAATTAACGTAACAAGCTGACAGGAGGCAATTCACAGTTCACCTCCACT 5864
Db 3490 GTTCTCTTAATTAACGTAACAAGCTGACAGGAGGCAATTCACAGTTCACCTCCACT 3549
Qy 5865 CCAAGACAGTGTGTTGAAGAGAGCACAACATGTGACTGCGGTAGAGCAGATTAAC 5924
Db 3550 CCAAGACAGTGTGTTGAAGAGAGCACAACATGTGACTGCGGTAGAGCAGATTAAC 3609
Qy 5925 TACATTTTACACATGAGCCCAAGCAAAATTTATGTTTGGCTATGCGGCAAGAGT 5984
Db 3610 TACATTTTACACATGAGCCCAAGCAAAATTTATGTTTGGCTATGCGGCAAGAGT 3669

5985 CCACCTGCAATGCTGAATGTAAACACCGCGCCGACACATAATTGGAGAACACATAAG 6044
3670 CCACCTGCAATGCTGAATGTAAACACCGCGCCGACACATAATTGGAGAACACATAAG 3729
6045 TCGACCAAGAATTCAGGCGGCGAGTTTCCAAACACTTTGGAACTGGCTGCTTGCAGTGT 6104
3730 TCGACCAAGAATTCAGGCGGCGAGTTTCCAAACACTTTGGAACTGGCTGCTTGCAGTGT 3789
6105 TTGGGGGAGCATCATCCCTCATTTGTTAGAGCTTAATAGTGTGTGCTGCGAGCTTATGC 6164
3790 TTGGGGGAGCATCATCCCTCATTTGTTAGAGCTTAATAGTGTGTGCTGCGAGCTTATGC 3849
6165 TTATTAACACAGTATGATGACTGAGCGGACACATGATAGCGGTAAACCTGCATGAC 6224
3850 TTATTAACACAGTATGATGACTGAGCGGACACATGATAGCGGTAAACCTGCATGAC 3909
6225 TTCCGAGGAGCGGTGTCATTAATGCCACGCGCGCTTGACACTAAACCTGCATGATTT 6284
3910 TTCCGAGGAGCGGTGTCATTAATGCCACGCGCGCTTGACACTAAACCTGCATGATTT 3969
6285 CCGAGGAGGACAGTGCATTAATGCTGTGAGTGCATTAATGCTGTATATCATCACTAAT 6344
3970 CCGAGGAGGACAGTGCATTAATGCTGTGAGTGCATTAATGCTGTATATCATCACTAAT 4029
6345 ATTAACACACTATATCATCTTTATGAGACTCATATGGTCTTAATATACATACACA 6404
4030 ATTAACACACTATATCATCTTTATGAGACTCATATGGTCTTAATATACATACACA 4089
6405 TATTTACTTAAAAACACTATACACTTATTAATCTTTTATTAATTTTCTTTTGT 6464
4090 TATTTACTTAAAAACACTATACACTTATTAATCTTTTATTAATTTTCTTTTGT 4149
6465 T 6465
4150 T 4150

RESULT 4

US-09-507-362-103
Sequence 103, Application US/09507362
Publication No. US20030096397A1

GENERAL INFORMATION:

APPLICANT: Dubensky Jr., Thomas W.

Belli, Barbara A.

Schlesinger, Sondra

Frylov, Sergey A.

Frylov, Ilya

TITLE OF INVENTION: RECOMBINANT ALPHAVIRUS-BASED VECTORS

WITH REDUCED INHIBITION OF CELLULAR MACRO-MOLECULAR

SYNTHESIS

NUMBER OF SEQUENCES: 125

CORRESPONDENCE ADDRESS:

ADDRESS: Seed Intellectual Property Law Group PLLC

STREET: 701 Fifth Avenue, Suite 6300

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/507,362

FILING DATE: 18-Feb-2000

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: McMaesters, David D.

REGISTRATION NUMBER: 33,963

REFERENCE/DOCKET NUMBER: 930049, 45706 /1196, 011

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 103:

SEQUENCE CHARACTERISTICS:

LENGTH: 11740 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 103:

US-09-507-362-103

Query Match 36.7%; Score 2380.4; DB 11; Length 11740;

Best Local Similarity 64.6%; Pred. No. 0;

Matches 3693; Conservative 0; Mismatches 1986; Indels 41; Gaps 8;

580 ACTGACGATGAAAGCGGAGCGTATATTTCTCATCGGAAACAGGCCAAGTCACTTC 639
5746 ACTGACGATGAAAGCGGAGCGTATATTTCTCATCGGAAACAGGCCAAGTCACTTC 5805
640 AACGAAATTCAGTACGTAATGTAACATCAACAAACCTATATTTGGATGGGCCCTGCATG 699
5806 AAAAGAAATTCAGTACGTAATGTAACATCAACAAACCTATATTTGGATGGGCCCTGCATG 5865
700 AAGATATTTAGCGCCCGCGCTCGATCGAAAGAGAAATGTTACAGAAAGAACTGC 759
5866 AAGATATTTAGCGCCCGCGCTCGATCGAAAGAGAAATGTTACAGAAAGAACTGC 5925
760 AATTATGCGCTCTGAAGAAATGAAAGAGATTCATCAAGAAAGTAAATATGA 819
5926 AAGATATTTAGCGCCCGCGCTCGATCGAAAGAGATTCATCAAGAAAGTAAATATGA 5985
820 AAGCAATTAACGCGGAGGAGTCAATTTCTGATTTGGGACATATCTATCAAGAAAGTGA 879
5986 AAGCAATTAACGCGGAGGAGTCAATTTCTGATTTGGGACATATCTATCAAGAAAGTGA 6042
880 ATCTGTGAGTGTATACAGATCAATTAATCTGTATCAATCTGCTCAACGCTAATTA 939
6043 ATCTGTGAGTGTATACAGATCAATTAATCTGTATCAATCTGCTCAACGCTAATTA 6102
940 ACAGTTTACATCTGACAGAGTGGCGGTTAAACGTGCACTTATGTTATCCAGAAATTT 999
6103 ACAGTTTACATCTGACAGAGTGGCGGTTAAACGTGCACTTATGTTATCCAGAAATTT 6162
1000 ACCCTACAGTACGAGTATTTGTAACAGATGAATGAGATGCGTATTTGACATGTTGG 1059
6163 ACCCTACAGTACGAGTATTTGTAACAGATGAATGAGATGCGTATTTGACATGTTGG 6222
1060 ACCGCGATCGTGTGCTGTATGATACAGCCACTTTTTCGCGCTAACTGAGAGCTAC 1119
6223 ACCGCGATCGTGTGCTGTATGATACAGCCACTTTTTCGCGCTAACTGAGAGCTAC 6282
1120 CAAGAAAGCATAGCTATTTTTCGAGCCAGATTAATGATGATGTTTCAAGAAATTCGAT 1179
6283 CAAGAAAGCATAGCTATTTTTCGAGCCAGATTAATGATGATGTTTCAAGAAATTCGAT 6342
1180 ATACATTACAAATGATATTTGCTGCTGACGATTAATTAAGAAATTCGATGATGATG 1239
6343 ATACATTACAAATGATATTTGCTGCTGACGATTAATTAAGAAATTCGATGATGATG 6402
1240 ATACATTACAAATGATATTTGCTGCTGACGATTAATTAAGAAATTCGATGATGATG 1299
6403 ATACATTACAAATGATATTTGCTGCTGACGATTAATTAAGAAATTCGATGATGATG 6462
1300 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1359
6463 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6522
1360 TTAGCAATATGTAACAAAGTGAAGAGGCGGAAAGCAGCAGCATTTGTTGCGAATATC 1419
6523 TTAGCAATATGTAACAAAGTGAAGAGGCGGAAAGCAGCAGCATTTGTTGCGAATATC 6582
1420 ATAAATCTAAACCGTTCGAGAGATTAACATGATGATGATGATGATGATGATGATGATG 1479

```
Db 6583 ATATATTTGGTCCCATTCGACAGAAAGTGCCTATGATAGATTGTGATGACATGAAAAGAG 6642
Qy 1480 ATGTCAAAAGTTACTCTCCGGCAGAAACATACAGAGAGCGGCTTAAGTGCAGGTTATTC 1539
Db 6643 ACGTGAAGATTACACAGAGCAGAAACACACAGAAAGAAAGACCCAAAGTACAAAGTGTATAC 6702
Qy 1540 AGGCTGAGATCCCTTGTCTACCGCTTACCTTGTGCGGAGTCCATCGGGAATTAAGTCCGTA 1599
Db 6703 AAGCCGAGAAACCCCTGGCAGCTCTTACTTATGCGGATTCACCGGAATTAAGTGTGTA 6762
Qy 1600 GACTGAATGCGGTCTTCTGCAAAATATCCACTCTCTTGACATGTCAAGCGGAAGATT 1659
Db 6763 GGGTTACGGCGCTGTGCTTCCAAATACATGACAGCTTTTGAATGTGCGCGAGGATT 6822
Qy 1660 TTGATGCGATTATTTGCTGAACATTTCCACACGCGGACCCAGTATTTGGAACGCAATCG 1719
Db 6823 TTGATGCAATCATATGACGAAACACTTCAAGCAAGGCGACCCGGTACTGAGACGGAATATCG 6882
Qy 1720 CGTGGTTTGAATAAAGGAGAGACGCTATCGCATTTGGCGGTGATGATCCCTGAGG 1779
Db 6883 CATATTGCAAAAAGCCAGAGACGCTATGGCGTTAACGGTCTGATGATCTTGAGAGG 6942
Qy 1780 ACTTAGGTGTCGACCAACCGCTCTTAGATTGTATGAGAGCGCGCTTGCAGAAATACAT 1839
Db 6943 ACCTGGGTGGATCAACACTACTGACTGTGATGCGATGGCGCTTTGGAGAAATATCAT 7002
Qy 1840 CTGTGCACCTACCTACAGAGAACGAGGTTTAAATTTGGTCCATGATGAATCCGGTATGT 1899
Db 7003 CCACCCATCTACCTACCGGGTACTCGTTTAAATTCGGGGCGATATGAATCCCGAAATGT 7062
Qy 1900 TCTTAAAGCTGTTGTGTCAACACACTAGTCAATATCATGATTTGTAGAGAGATCTAGGTG 1959
Db 7063 TCTTCAACATTTTGTGTCAACACAGTTTGAATGTCGTTATGCGCAGAGAGTACTAGAG 7122
Qy 1960 AACGGTTAACCAACGTCAGCGTGCAGCGCTCTATCGCGAGATGATGATAGTATGATGATG 2019
Db 7123 AGGGGCTTAAAGCTCCAGATGTGCAAGGTTCAITTTGGAGACGACATCATATCATGAGAG 7182
Qy 2020 TCGTCTCCGACACTTGAATGGCGGAGAGATGCGCACCTTGGCTGAACATGGAATMAAAA 2079
Db 7183 TAGATCTGACAAAGAAATGCTGAGAGGTGCGCACCTGCTCAACATGAGAGGTATAGA 7242
Qy 2080 TTATTTGAGATGATTTGATGATCAAAAGACCTCTCTGTGCGGAGTTTATCCTGATGG 2139
Db 7243 TCATGCGAGATGATGATGATGAGACCACTTACTTCTGCGCGGATTTATCTTGCAAG 7302
Qy 2140 ACCAGATTAACAGGACGCTGACAGATCGCAGACCTCTAATAAAGCTTTTAAAGCTTG 2199
Db 7303 ATTGGTTACTTCCACAGCGTGCAGCGTGGCGATCCCTGAAAAGCTGTTTAAAGTTGG 7362
Qy 2200 GAAAAACATTGCGCATGATGATACCAAGATCGACCGCGCGGCACTGATGATG 2259
Db 7363 GTAAACCGCTCCCGCGAGCAGACAGAGCAAGAACAGAAAGCGGCTCTGTCTAGATG 7422
Qy 2260 AAGCAATGCGATGGAACAGATTGGAATTAACGAGATTTAGTAGAGCGGTAGAAATCCA 2319
Db 7423 AAACAAAGGCGTGTGTTTAAAGTAAAGTAAACAGCACTTTAGCAGTGGCCGTGACAGACC 7482
Qy 2320 GATACGAGATCATATGCGAGGCGCTGATCATCATGCTCTGCTCCACGTTAGCCGAAAGCG 2379
Db 7483 GGTATGAGTGAACAATATTAACACTGTCTTACTGAGATTTAGAAATTTTGGCCAGAGCA 7542
Qy 2380 TTAAAGACTTCAGAGACATAAGAGGAGCCCATCATCCCTTATAGCTGACCTTAAATAGG 2439
Db 7543 AAAGAGCATTCAGAGCATGAGAGGGAATAAAGCAATCTCTACGATGCTCTTAAATAGT 7602
Qy 2440 TGAAGTATGACAGCGACCTAC-----CCACCGGAGAGATGTTTCCATACC 2486
Db 7603 CAGCATATGATCTTCACTGACTAATACTAACAAACACACCACTGATGATGAGAGATTTCT 7662
Qy 2487 CTCAGCTGAACCTTCCACAGTTTACCTTACAAATCCGATGGCTTACGAGATCCAAACC 2546
Db 7663 TTAACATGCTCGGCGCGCCCTTCCCGGCCCCCACTGCCATGTGAGGCGCGAGAA 7722
Qy 2547 CTCCTAGCGCCCGCTGAGAGCGGTTTCGGCCCCCGCTGGCTGCTCAATGGAAGATCTTA 2606
Db 7723 GAGAGCAGCGCGCCCGATGCTGCGCGCAACGGGCTGGCTTCAATCAGCACTGTA 7782
Qy 2607 GAGAGTGCATATGATCACTTGAATTTCAACAGATCATCTTAAATCCGCGCCAGTCCA- 2665
Db 7783 CACAGCGGTCAATGCGCTGATGATGAGACAGGCACTAGACTCAACCCCAAGTCCAC 7842
Qy 2666 -----CCGCAAAAGAAAGAAAGATGCTCTTAAAGCCAAACCTTCTACCTTAAAGA 2720
Db 7843 GCCCGCACCGCGCCAGAAAGACAGCGCCCAAGCAACCAAGGAGCCAGAAACCAA 7902
Qy 2721 AGAAGCAGCAAGCCAGAGAGCAAGAAC-----GCAAGCTTAAACAGGGAACGACCACTGA 2777
Db 7903 AAACGACGAGAAAGAAAGAAAGCAACCTGCAAAACCCAAACCCGGAAGAGACAGGCA 7962
Qy 2778 TGTGTATGAAGTTGAGTGCAGACAGACATTTCCGATCATG-----CTGAAAGGCCAAGTGA 2834
Db 7963 TGGCATTAGTTGAGAGCCGACAGATTTGTCAGCTCAAGAACGAGAGAGAGATGTCA 8022
Qy 2835 ATGATATGCTGCTGCTGTCGAGAGAGCTGATGAACCACTTCCACTTGAAGAAAAA 2894
Db 8023 TCGGGCAGCAGCTGGCCATGGAAGAAAGGTAATGAACCTCTGCACTGTAAGAAACCA 8082
Qy 2895 TTGATATAGCAATTAAGCCGCGCTGAAATTTGAAGAGGTACATGATAGACTGTGAGT 2954
Db 8083 TCGACCACTGTGCTATCAAACTCAAAATTTACCAAGTGTGACATGACATGACATGAGT 8142
Qy 2955 ACGGCGAGTTCCTCCAGAACATGAAATCAGACACGCTGAGTACACAGCGACAAACAC 3014
Db 8143 TCCGACAGTTGCGAGTCAATGAGAGATGAGGCAATTCACCTACACAGTGAACACCCCG 8202
Qy 3015 CGGGCTTCTAACACTGGCACACCGGCGAGTCCAATGTGAATGGAATTTAACCTGATC 3074
Db 8203 AAGGATTTCTAATCTGGCACACGAGCGGTGAGTATGAGAGTGAATTTTAAACATCC 8262
Qy 3075 CGAGAGAGTGGCGGGGAAAGGAGACAGCGGAAAGACCGATCCGTGACAAAGAGGCGAG 3134
Db 8263 CTCGGAGTAGAGAGGAGAGAGACAGCGGTGCTCGATCATGATGATATCTCGGTGCG 8322
Qy 3135 TTGTGCTATTTGTTCTAGAGAGTGCAATGAGGACGCGCTTACGGCGCTTTCAGTGTCA 3194
Db 8323 TTGTGCGGATGATCTCTGCTGCGGTGATGAAAGAACAGCACTGCCCTTTCGCTGTCA 8382
Qy 3195 CTGGAACCAAGAAAGGGGTACCAATTAAGGATACCCCGAAGTTTGAAACCTGTG--- 3251
Db 8383 CCTGGAATAGTAAAGGGAACCAATTAAGACGACCCCGGAGAGGACAGAAAGTGTGCG 8442
Qy 3252 -----CACTAGTTACAGCGCTATGGGTGCTTGCATGTCAAGTCAAGTCCCATGCGACAAAC 3305
Db 8443 CAGACCACTGTGTCAGGCAATGTGTTGCTCGAAATGAGATCTCCATGCGACCGCC 8502
Qy 3306 CACCGGTGCTATTTCACTGACCCAGAACGAAACACTGCACTGTGCTCGAAGAAACGTG 3365
Db 8503 CGCCCAATGCTATACCGCGGAACCTTCCAGACCCCTCGACATCTTTGAAGAGACGTGA 8562
Qy 3366 ACAATCCAAATTAACGACACGCTCTGAGAGACGTCTTGAATGTCCATCAC---GCCGCG 3422
Db 8563 ACCATGAGCGCTACGATACCTGCTCAATGCAATATGCGGTGCGGTCTCTGCGAGAA 8622
Qy 3423 CCAAGCAAGCAATTAACGATGATTTCACTGACCTGACAGTCCCTGAGGGTTGCGCGCT 3482
Db 8623 GCAAAAGAACGTCATTTAGACCTTACCTGACAGCCCTTACTTGGGACATGCTGCT 8682
Qy 3483 ATTGACAGCACTCAACGCGTGTTTGACGCCAATAAATTTGAACGCTGTGGAGCAAT 3542
Db 8683 ACTGCCACCACTATGTAACCGTGTCTTCAAGCTTGAAGATGAGAGGTCTGGGAGCAAG 8742
Qy 3543 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3602
Db 8743 CGAGCGATACCACTATGCGATACAGACTTCCCGCAGTTTGGATGATGACCAAGCGGAG 8802
```

QY 3603 CTGGGATGTCCAAATTCGGTTACATGCTTTTCGACCAGACCATGATCAAGAG 3662
DB 8803 CAGCAAGGGCAAAAGTACCGCTACATGCTTAAACAGAGTACACCGTTAAAGAG 8862
QY 3663 ACAGATGTGAGAAATAGCTATCAGACATCTGCACCTCGCTGCTTGGCCCAAG 3722
DB 8863 GCACCATGAGATCAATCAAGATTAGCACCTCAGGACCGGTGTAAGGCTTGTCAAG 8922
QY 3723 GGATCTCTCTGTGACTCAATGCTCTCGAGGACAGTGAACGCTGATTCACAGACG 3782
DB 8923 GATCTCTCTCTCGCAAAATGCGCTCCAGGGAGAGGTAAACGTTAGCATGAGATA 8982
QY 3783 GAGCATCTGAAATTCATGACACCGTGGAGAAATAGATCAGAGAGATTTGCTGATAG 3842
DB 8983 GCACTCAGCAACGTGATGATCACTGCGCCGCAAGTAAACCAAAATTCGTGGACGG 9042
QY 3843 AGGATCTTGTCTCCACCGCTCCATGGAAAGCTGGTAAAGTCCACGTTTACGATCACT 3902
DB 9043 AAAATATGATCTACCTCCGTTACCGTTAAATAATCTTTCACAGTGTACGACGTC 9102
QY 3903 TGAAGGACGCTCGCGGGTACATTAACATGACAGGCGCCACAGCGGTATAGT 3962
DB 9103 TGAAGGAAACAATGCAAGGCTACATCTATGACAGGCGGAGACCGCATATATCAT 9162
QY 3963 CCTATCTGAGGAGCGCTCAGGCGAAGTGTACATTAACACACTTCTGGCAAGACGTC 4022
DB 9163 CCTACCTGGAAGATTCATCAGGAAAGTTTACGAAACCGCCACTGTGGAAAGAACATTA 9222
QY 4023 CTTACGAATGTAGTGTGCGCATTCACACAGGATTCGTAGCACCGGAAAGAAATGA 4082
DB 9223 CGTATGAGTGCAGTGTGCGGCATCAAGACCGGAACGTTTTCGACCGGCAACCAATCA 9282
QY 4083 ACGGCTGACCTAAAGCAAAACAGTGCATTTGCTCAAGAGGACCAACGAAATGGGCT 4142
DB 9283 CTGCTTGACCGCCATCAGAGAGTGCCTGCTTAAAGAGGACCAACGAAAGTGGCT 9342
QY 4143 TCAACTCGCGGATCTTATAGGACACAGACCACTCAGTGAAGTAAATTTGCATTC 4202
DB 9343 TCAACTCAGCGGATCTTATAGGACACAGACCACTCAGTGAAGTAAATTTGCATTC 9402
QY 4203 CATTCGCTTGACACCGAGCTGTGCGGCTTCCGTTAGCTACACGCTTACAGTCA 4262
DB 9403 CTTTCAAGTTGATCCCGAGTACCTGATGCTCCGTTGGCCAGCGCGCATGTAAATAC 9462
QY 4263 AGTGTCTCAAGGATCACTCTCCACCTGACCTGACATTCGACCAATTTGCTGACACGA 4322
DB 9463 ATGGCTTTAAACATATCAGCTCCCAATTAGATACAGACCACTTGTGCTCACACCA 9522
QY 4323 GAAATTTGGGCTCGAGCAGACAGCAGATGATTAACAGGCTCTACATCCAGGA 4382
DB 9523 GGAATCTAGGGGCAAAACCGGAAACCACTGATGATTCGTGGAAGACCGGTCAAA 9582
QY 4383 ATTTTCTGTGGGCGAGAAAGGCTGAGTACGTATGGGTTAACCATGAAACAGTCAAG 4442
DB 9583 ACTTCAACGCTGACCGAGATGGCTGGAAATACATATGGGAAATCATAGGCAATGAGG 9642
QY 4443 TCTGGGCGCAGAGTCTGGACACAGGCGAACCATGATGATGATCATCATCC 4502
DB 9643 TCTATGCGCCAGAGTCTGACACAGAGACCTTCACAGATGCGCACACCAAAATAGTAC 9702
QY 4503 ACTATATCATCGGATCACTGATCACTGTCAATTTGCTGTGTGTGCTGTGCTA 4562
DB 9703 ATTACTACATCGCATCTCTGTATACATCTTAAAGCGCTGCAATCACTACCGTGGCG 9762
QY 4563 TCTGTAGGAGCATGATCATCAGACGCTTGCATTCGCAAGCAAGAGAGATCGCTGA 4622
DB 9763 TGATGATGCGGTAACTGTGCAAGTATATGCTGTAAAGCGGCGGTGAGGCGCTGA 9822
QY 4623 CGCCATACGCGCTTGCACCGAAACGAGTACCAAGCATTTAGCGGTTTGTGTGCA 4682
DB 9823 CGCCATACGCGCTTGCACCGAAACGCGTAAATCCCAATTCGCTGCACTTGTGTGCG 9882

QY 4683 TTCCGCCAACCAACGCTGAAACATTTGGAGAACTTTGAACCATCTGTGTTTAAACAAC 4742
DB 9883 TTAGTGGGCCAAATGCTGAAACGTTCAACGAGACATGAGTTACTGTGTGCAACATC 9942
QY 4743 AACGTTTCTGTGGGACAGTTGTGATCTCTCTGGACGCGCTGTTATTTCTGTTCCCT 4802
DB 9943 AGCGTTCTTGTGGGTCAGTGTGATCACTTTGGCGCTTTCATCTTCAATGCGCT 10002
QY 4803 GCTTTATGCTGATGCTTTTATTTGTTGAGAGGCTGTGCTGGGAAAGTAGACG 4862
DB 10003 GCTGCTCTGCTGCTGCTTTTATTTAGTGTGCGGCGCTTACCTGGCAGAGTAGACG 10062
QY 4863 CTTTGAACATGCGACCATGCTGCAATGTTCCGGGATCCGTTAAGCGCTGTGCTG 4922
DB 10063 CTTAGCAACATGCAACCATGCTTCCAAATGTGCCAAGATACCGTATTAAGCACTTGTG 10122
QY 4923 AACGCGAGGTTACGCGCACTTAACCTGAGATCAAGGTGCTCATGGAAATTAACAC 4982
DB 10123 AAAGGCGAGGTTAGCCCGCTCAATTTGGAGATCACTGTATGCTCTCGAGGTTTGC 10182
QY 4983 CTTCACTAACAGAGATGAGTACCTGCAAAATTCACACAGTATCTTACACACAG 5042
DB 10183 CTTCACTAACAGAGATGAGTACCTGCAAAATTCACACAGTATCTTACACACAG 10242
QY 5043 TTAATGCTGCGGCTCCCTCAGTGTCAAGGCAATCTCAAGGCGGATTAACATGCCGCG 5102
DB 10243 TCAATGCTGCGGCTCCCTGGAATGTGACGCGCGCGCTCATGACATATCTGTCAAG 10302
QY 5103 TTTTGGCGGTGTACCTTTCATGTGGGAGCGCAATGCTTCTGTGACAGTGA 5162
DB 10303 TCTTGGAGGGGTCTACCCCTTTATGTGGGAGAGGCGCAATGTTTTCGACAGTGA 10362
QY 5163 ACAACAAGTGTGAGGCGGAGTGTGAGTTGCTCAGATGATGATATGATACAGGAG 5222
DB 10363 ACAGCCAGATGAGGCGGAGTGTGAGTTGATGACAGATGAGTGTGCTGTACACGCGC 10422
QY 5223 TCGCACTAAAGATTCAACAGCTCTCTGAAAGTGTGCGCTGCTATATAGGCAACA 5282
DB 10423 AGCGGATTAAGTGTACATCTGCGCGCATGAAAGTATGAGATCTGTATGTGTACCGGA 10482
QY 5283 CCACCGGCACTGTGATACGTTTGTCAATGTGCGGTACGCGCAGGTTCTCAGCGGACCTGA 5342
DB 10483 CTACCAAGTTTCTGTAGATGTGATCGTGAACGAGTCAACACGAAACGCTTAAAGCTTGA 10542
QY 5343 AGGTATAGCAGGCGCATATCAGCCGCTTTTTCACCTTGTGACATTAAGTCTGTATCA 5402
DB 10543 AAGTATAGTGTGACCAATTTCAACATCTTTTACGCAATTCGATCATTAAGTCTGTATCC 10602
QY 5403 GAAAGGGGCTTGTTAACAACATTAACGATCTTCCGTGATGTGAGTATGAAACAGAGGCT 5462
DB 10603 ATGCGGCGCTGTGTACATCTATGATCTTCCGGAATATGAGCGATGAAACAGAGGCT 10662
QY 5463 TCGGCAATTTCAAGCATCTCTGCTGATGTATGATGATGATGATGATGATGATGATG 5522
DB 10663 TTGAGACATTTCAAGTACCTCTTGTAGTACAGAGATTCATGCGGACAGACATTA 10722
QY 5523 GGTCTGTGAAGCTTCTGTGTAAGAAATCAACGCTTCCCTTACACCAAGCAATATCAGGCT 5582
DB 10723 GGTCTGTGAAGCTTCTGTGTAAGAAATCAACGCTTCCCTTACACCAAGCAATATCAGGCT 10782
QY 5583 ATGAAATGTGGAAGAAACAATCAGAGAGACCCCTGCAAGAAACAGACCAATTTGATGAT 5642
DB 10783 TTGAGATGTGGAAGAAACAATCAGAGAGACCCCTGCAAGAAACAGACCAATTTGATGAT 10842
QY 5643 AAATTTGAAGTGAAGCTCTGCGAGCGTCTAATCTGTGCTTACGGGACATCCCTATCTGA 5702
DB 10843 AGATTTGACATTAATCCGCTCGAGCGGCTGAGCTGTTCATATAGGGAAATTTCCATTTCTA 10902
QY 5703 TTGACATCTCTGATGAGCTTTTGTGATCATATGATATCAACCAATTTTAAAGTGA 5762
DB 10903 TTGACATCTCTGATGAGCTTTTGTGATCATATGATATCAACCAATTTTAAAGTGA 10962
QY 5763 GCTGCAAGTATGACAGATGATTAATCTGACAGATTTGTGTGTTCTTAACATTAAGT 5822

Db 10963 AATGTGAAGTCAGTGAAGTCACTATTACAGAGCTTCGGGGAGTGGCCACCTCGAGT 11022
QY 5823 ACAAGCTGACAGGAGGAGGACATTGCGAGTTCCTCCACACAGAGCTGTTTGA 5882
Db 11023 ATGATCCGACCGGAGGTCATGAGCCCGTACATTCGACATTCGAGACAGCACTCTCC 11082
QY 5883 AGGAGGACACACAGTGAAGTCTGCGTAGGACACATTAACATTAATTTAGACATCGA 5942
Db 11083 AAGGTGACAGTACATGCTCTGAGAAAGAGCGGTGACAGTACCTTTAGCAGCCGGA 11142
QY 5943 GCCCACAAGCAAAATTTATAGTTTCGCTATGCGGCAAGATCCACCTGCAATGCTGAT 6002
Db 11143 GTCCACAGGGGAATTTATGATGATGCTGTGTGGAGAGAGACAATGCAATGACAAAT 11202
QY 6003 GTAAACACAGCGGCGGACCACTAATTTGAGAACACATTAAGTGCACCAAGATTCAGG 6062
Db 11203 GTAAACACAGCGTACCATATCTGTAGCACCCCGCACAAATGACCAAGAAATTTCAAG 11262
QY 6063 CGGAGTTTCCAAACATCTTGAACCTGGCTGTGACCTGTGGGGAGCATCATCCC 6122
Db 11263 CCGGATCTCAAAAACATCAGAGGCTGGCTGTGCTTTGCGGCGGCTCTGTCGC 11322
QY 6123 TCAATTTGTAGAGACTTATAGTGTGCTGCTGACCTCTATGCTTAAACACAGTAGAT 6182
Db 11323 TATTATTAATAGAGCTTATGATTTTGTCTGACACATGATGCTGACAGCAGAAAT 11382
QY 6183 GACTGACGGGAGACACTGA-CATAGCGGTAAACTCGATGATCTTCGAGAGACGTGCT 6241
Db 11383 GACCGCTACGCGCCCAATGATCCGACAGCAAAACTGATGATCTTCGAGAGACGTGAT 11442
QY 6242 GCATATGCGACGCGCGCT 6261
Db 11443 GCATATGCTACGCTGCT 11462

RESULT 5

US-09-901-106-1

Sequence 1, Application US/09901106
Patent No. US20020151067A1

GENERAL INFORMATION:

APPLICANT: Garoff, Henrik

Liljestrom, Peter

TITLE OF INVENTION: DNA Expression Systems Based on

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Birch, Stewart, Kolasch & Birch

STREET: P.O. Box 747

CITY: Falls Church

STATE: Virginia

COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/901,106

FILING DATE: 10-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/920,281C

FILING DATE: 13-AUG-1992

ATTORNEY/AGENT INFORMATION:

NAME: Murphy Jr., Gerald M.

REGISTRATION NUMBER: 28,977

REFERENCE/DOCKET NUMBER: 828-103P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-241-1300

TELEFAX: 703-241-2848

TELEX: 248345

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 11517 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: RNA (genomic)

HYPOTHEetical: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Semliki Forest Virus

FEATURE:

NAME/KEY:

LOCATION: 1..11517

OTHER INFORMATION: /label= genome

/note= "Semliki Forest Virus complete nucleotide

sequence, presented as a cloned DNA sequence; see

Figure 5."

FEATURE:

NAME/KEY: CDS

LOCATION: 87..7379

OTHER INFORMATION: /product= "SFV polyprotein"

FEATURE:

NAME/KEY: CDS

LOCATION: 7421..11179

OTHER INFORMATION: /product= "SFV polyprotein"

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-901-106-1

Query Match 22.24; Score 1442.6; DB 10; Length 11517;

Best Local Similarity 55.54; Pred. No. 0;

Matches 3130; Conservative 0; Mismatches 2409; Indels 102; Gaps 14;

QY 594 GCGGAGCGTATATTTCTGATCGGAAACAGGCCAGGTCACTTCAACGAATTCAGTA 653
Db 5529 GCGGAGCGTATATTTCTGATCGGAAACAGGCCAGGTCACTTCAACGAATTCAGTA 5588
QY 654 CGTCAATGTAACTACAGAAAGTATTTGATCGGCGCTCCATGAGATTTACGCC 713
Db 5589 AGGAGACATCTCCAGTCCGACACATGATGCGTCCAGAGAGAAATGATACCCG 5648
QY 714 CCGGCGCTCGATCTCGAAGAGAAATGTTACAGAAAGTCAATTTATGCGCTCT 773
Db 5649 CCAAAATGATATCTGAGAGAGAGAGCTGTTCTGTTGAAATGAGATGACCATCG 5708
QY 774 GAAGGAATGAAAGCAGTATCAATCAAGAAAGTAAATTTGAAAGCAATTACAGCG 833
Db 5709 GAGGCTAATAGAGTCCATACCAAGTTCGCAAAAGTGAAGACATGAAGCAGCGTGTG 5768
QY 834 GAGGAGCTATTTCTGATTTGGGCAATATCTATCATCAGAAAGTAACTCTGCAAGTGT 893
Db 5769 GACAGGCTACATCGGGGCGCAATTTGTACAGGAGCGAGCTGAGCCGATACCAACA 5828
QY 894 TACAGATCAATTAATCTGTACCAATCTCTGTCACAGGTAATTTACAGGTTTACATCT 953
Db 5829 TAGCGGTTTGGTAAACCCCGCCCGGTACTCCCTTACGTAATGAAATTTCTCAAGC 5888
QY 954 GCAGAGTCCGGTTAAAGCTGCACTTATCTTATCCAAAGAAATTTACCTTACAGTGGCC 1013
Db 5889 CCGGATGTACCAATCGCAGCGTGCAGAAATACCTTATCCAAATTTACCAACAGTGGCG 5948
QY 1014 AGTATTTGTAATACAGATGAATAGATGCGTATCTTACATGAGTGGAGCGGATCGTGC 1073
Db 5949 TCGTACCAATTAACAGATGAATAGATGCGTATCTTACATGAGTGGAGCGGATCGTGC 6008
QY 1074 TGTCTAGATACAGCACTTTTGTCCGGCTAAATGGAAGTAACTCCAAAGAAAGCATAGC 1133
Db 6009 TGTCTAGATACAGCACTTTTGTCCGGCTAAATGGAAGTAACTCCAAAGAAAGCATAGC 6068
QY 1134 TATTGAGCGCAAGATTAAGTACAGCGGTCCATGCGCTTATACGAATTAATTAATTAAT 1193
Db 6069 TACCACAGCGCACTGACGAGGCGGTCCGCTCCGCTTTCAGAAACACACTACAGAAC 6128

Qy 1194 GTATTGCTGCACTACTATTAAGAAATTGCAACGTTACCCAAATGCGAATTAACCTGTC 1253
Db 6129 GTGTAGGCGCGCCACCAAGAAAGAACTGCAACGTCAGCAAAATGCGAATTAACCCACC 6188
Qy 1254 TTAGATTGGGGGCGCTTTAATGTTGATTGTTCAAGAAATACGCAATGATAGTAC 1313
Db 6189 ATGAGCTGGGAGTGTTCACGTTGAGTGTTCAGGCGTATGCTTCCGCAAAATAT 6248
Qy 1314 TGGGATACCTTTCGGATTAACCTTATGCGGCTAATACAGAGAGGTTAACGCAATATG 1373
Db 6249 TGGGAAGATATGCTTAACAACTTATCCGATTAACCACTGAGAACATCACTATG 6308
Qy 1374 ACAAGCTGAAGGGCCGAAAGACAGACATGTTGGCAATATCATTAATTAAGACG 1433
Db 6309 ACCAAATGAAGGCCGAAAGCTGCTGCTTGTTCGTAAGACCAACTGTTCCG 6368
Qy 1434 TTGCGAGGATACCAATGATCAATTCGTCATGATCTTAAGAGATGTCAAAGTTACT 1493
Db 6369 CTGCGAGGAGTTCCTCATGAGACGATTCACGCTCGACATGAAACGAGATGTCAAAGTCACT 6428
Qy 1494 CCGGCGAGAAACATACAGAGAGGGGCTAAGGCTTAAGGCTTATTCAGGCTGACATGCC 1553
Db 6429 CAGGAGGAAACACAGAGAGAAAGCCAAAGTCAGGTAATTAAGAGCGAGCA 6488
Qy 1554 CTGCTACCGCTTACCTTTCGCGGATCCATCGGGAATTAGTCCGTAAGCTGATCGG 1613
Db 6489 TTGGGACCGCTTACTGTGGGATCCACAGGAAATTAAGAGATTAATCTGTG 6548
Qy 1614 CTTCGCAAAATATCATCTCTCTTCGACATGTCAGCGGAAAGATTTGATGCAATTT 1673
Db 6549 TTACCCCTTAACGTCACACATGTTGATATGTCGCGAAGACTTTGACGCGATCATC 6608
Qy 1674 GGTGAACATTTCCACAGCGGCGCCAGTATGGAACAGCATGCGCTGTTATATA 1733
Db 6609 GCTCTCACTTCCACAGAGAACCCGCTTACAGAGCGACATTCATTCACAAA 6668
Qy 1734 AGCGAAGACGACGCTATCGCCATTTTCGCGTGTGATGATCTTGAAGACTAGGTGAC 1793
Db 6669 AGCGAGACGACTCCTTGCTTACAGGTTAATGATCTCGAAGATCTAGGGTGGAT 6728
Qy 1794 CAACCGCTTATGATTTGATAGAGCGGCGTTCCGCAATATCATCTGTGACCTTACT 1853
Db 6729 CAGTACTGTGAGCTTATCGAGCGACGCTTTGGGAAATATCCAGCTGTCACCTACA 6788
Qy 1854 ACAGGAAGAGGTTAAATTTGTCGCCATGATGAATCCGATATTTCTTAACGCTGTT 1913
Db 6789 ACTGCGACGCTTCAAGTTGAGAGCTATGATGAATCGGCGATGTTCTGCTTGT 6848
Qy 1914 GTCAACACACTAGTCAATATCATGATGCTAGCAGAGTACTAGTGAACGTTAACAG 1973
Db 6849 ATTAACTGTTTGAACATCACTAGCAGAGGATCTGAGACAGACTCACTGAC 6908
Qy 1974 TCAAGTGCAGGCTTATGCGGACGATTAACATAGTATGATGTCGTCGACAC 2033
Db 6909 TCCGCTGTGCGGCTTATCGGCGACGACAACTGTTACGAGTATCTCCAGAAC 6968
Qy 2034 TTGATGGGAGAGATGCGGCACTTGGCTGAACATGAGAGTAAATTAATGATGCAAT 2093
Db 6969 CTGATGGGAGAGTGCCTGCTGTTGAGCAATGAGGTTAAATCATTAAGCTCTGC 7028
Qy 2094 ATTGATCAAAAGCACCCTACTTCTGTGGGGGATTATCTGCTGAGACAGATACAGG 2153
Db 7029 ATGGCGAAGAAACCCCATATTTTGTGGGGATTCATATGTTTGAACAGCTCACAG 7088
Qy 2154 ACAAGCTGACAGTGCAGACCTCTTAAGAGGCTTTTAAGCTTGAAGAACATATGCA 2213
Db 7089 ACCGCTGCGGTTTCAAGACCACTTAAGCGCTGTTCAAGTGGATAGCGCTAAC 7148
Qy 2214 GTGATGATACCAAGACTGCGACCGCGCGGCACTGATGATGAAGCAATGCGATG 2273
Db 7149 GCTGAAGCAAGACGAGACAGGCGACGAGACCTGATGAGAGGTTAGCAAGTGG 7208
Qy 2274 AACGAATTTGAATTAAGGAGATGTAAGGCGCTGATGATTCAGATACGATCATTA 2333

Db 7209 TTCCGAGAGGCTTGGGGCGGAAGTGAAGGCGACTTAACATCTAGATATGATAGAG 7268
Qy 2334 CTGGAGGCGCTGATCATACGCTCTCTCTGACAGTTAGCGGAAGGTTAAGAACTTCAAG 2393
Db 7269 GCGTCAAAAGATCTCTATAGCCATGACACCTTGGAGAGGACATTAAGCGCTTTAAG 7328
Qy 2394 AGCATTAAGAGGAGCGCAATCAACCTTACGCGCTGACCTAAATAGTACAGTATAGCA 2453
Db 7329 AAATTAAGAGAGCTGTTATACACTCTTACGCGGCTCTTAATGTTGCTGTTAATAC 7388
Qy 2454 CGACCTTACCAACCGGAGAAATGTTTCATACCTCAGCTGAATTTTCACCAAGTTTAC 2513
Db 7389 GAATCTGATATATAGCGCACTATTATAGCAACATGAATTAATCTCCCTACGCAAGCTTT 7448
Qy 2514 CTACAAATCCGATGCTTACCGAGATCCAAACCTCTCT-----AGCGCGGCTG 2563
Db 7449 ACGGCGCGGTGGCGCGCGCGCGCGCGCGCTTGGCGCTTGGCAGGCTCACTCCG 7508
Qy 2564 AGGCGGTTTGGCGCGCG-----CTGGCTGCTCAATGGAATCTTGAAGAGTGAATAG 2618
Db 7509 TGGCTCCGCTGCTCCGACTTCCAGGCGCGAGATGACGAACTCATGAGCGCTTA 7568
Qy 2619 TCAACTGACTTTCAACCAAGATCACTTAATCCGCGCAGGTCACCGCCAAAGAA 2678
Db 7569 ATGGCTGACATGAGACAGAAACGAATGCTCTGCTGATAGGCTCCAAACCAAGAA 7628
Qy 2679 AGAAGAGTCTCTTAAGCCAAACCTTACTACGCTTAAAGAAAGACAGCAAGCAAG- 2737
Db 7629 AGAAGACCAACCAACCAAGCCGAAACGAGCCGAAAGATCAACGAAACAGCAGC 7688
Qy 2738 -----AGAGGAAACGAGCG-----TAAACGAGGAAACGAGCAC 2774
Db 7689 AGCAAAAGAAAGAAAGACAGAGGCTGACAAAGAAAGAAAGAACCGAAAGAGAA 7748
Qy 2775 GTATGTATGATGAGTGGAGTGCAGACAGATTTCCGATGATGCTGAACGCGCAAGTGA 2834
Db 7749 GAATGTGATGAAGTTGAATGATGCTGATCTTGGAAGTCAACAGAGAAAGTGA 7808
Qy 2835 ATGATATGCTGCTGCTTGTGAGAGAGGCTGATGAACCACTCCAGTTGAAGAA 2894
Db 7809 CTGGTACGCTGCTGCTGCTGCGCAAAAGTATGAACTGCGCCACTGAAAGAGTCA 7868
Qy 2895 TTGATTAATGACATTAAGCGGCGCGGAAATGAAGAGGCTAGCAATGACATTTGAGT 2954
Db 7869 TCGACAGCGGACCTGCGAAAGCTAGCTTCAAGAAATCGAGCAAGTATGCTTGA 7928
Qy 2955 ACGGCAAGCTTCCGAGAACTGAATGACAGCGCTCAGTACACAGCGCAAAACAC 3014
Db 7929 GTGCCAATACAGATTCATGAGTGCATGCTCAAGAGTACACGATGAGAGCCG 7988
Qy 3015 CGGCTTCACTGACGACACAGCGCGCTGCTCAATAGAGATGAGATTTACCGTAC 3074
Db 7989 AGGCACTATTAACGTCACACCGGCGTGTTCAGTACAGCGAGTGAAGTTCATTA 8048
Qy 3075 CGAGAGAGTGGGCGGAAAGCGGACAGCGGAAACGATCTGACAAACAGAGGAGAG 3134
Db 8049 CGACAGGCGGCGCAACCGGAGACAGTGCAGCGCCATCTTGAACAAAGGAGAGAG 8108
Qy 3135 TTGTGCTATTTGTTCTAGAGAGTGCATTAATGAGGCGACGCTTACGCGCTTCA 3194
Db 8109 TAGTGTGATGCTCTGCGCGGCGCAACGAGGCTACGACAGCACTGTGCTGATCA 8168
Qy 3195 CTGGAACCAAGAGGCTGACATTAAGAGTATCCCGGAAAGTTTGAACCGTGT-- 3251
Db 8169 CTTGGAACCAAGATATGATGAC---TAGATGACCCCGAGAGGCTCCGAAGAGTGTCCG 8225
Qy 3252 ---CACTAGTAAAGCGCTATGCGCTTGTGAGATGTCACGTTCCCAAGCAAAAC-- 3306
Db 8226 CCGCGCTGATTAAGCAATGCTGTCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTG 8285
Qy 3307 -----ACCGGTGCTATTAAGTGAAGCGGAGAAAGCAAGTGAAGTGAAGTGAAG 3359


```
D8 8286 CGTGTACCTTGTGCTATGAAAAAACAACGAGAGCCACACTACGATGCTCGAGATA 8345
QY 3360 ACCGTCACATCCAAATTAACACACGCTGCTGGAGAAAGTTGAATGTCCA---TCAC 3416
Db 8346 ACGGTGATAGCCGAGGGTAATCAACGACTCTTCAGGACGCTTGAACGCGGAAACGGA 8405
QY 3417 GCCGGCCCAAACGAAAGCATTAACGATGACTCAGACAGT-----ACGATCCCTAAC 3467
Db 8406 CAGACACCGGCGGAGCGGTGTCGACACATTCACAGGTATTAAGCTACACGCGCTTACA 8465
QY 3468 TGGGGTCTGCGCGTATTCAGACACTCAACGCGGTGTTTCAGCCCAATAAAATTGAGA 3527
Db 8466 TCGGCTACTGCGCGCTGCGAGACAGGCACTGCTGTCATAGCCCGTACCAATTAAG 8525
QY 3528 ACGGTGGAGCAATCTGATGATGATGATGATTAATTCAGAGTCTCGGCAAAATTCGGCT 3587
Db 8526 CGGTACAGTCCGAAGCTACACGAGGATGCTGAAGATTCAGTTTCGGCAAAATTTGCA 8585
QY 3588 ACAATCAGGAGGAGTACGCGATGTACCAAAATTCGTTACATGCTTTTCAGCCAGCAAC 3647
Db 8586 TAGATTAAGAGTACATATGATCTACACAGAAATTAAGTA-----CGCAACCGGC 8636
QY 3648 ATGACATCAAGAGAGACAGATGAGAAAATGCTATTCAGCATCTGCACTCGCCGTC 3707
Db 8637 ACGCATTTAGAAATGCCGTCCGTCATCTTGAAGTAGCCACTCCGAGACATGTTTCG 8696
QY 3708 GTCTTGGCCACAAGGGTACTCTGTAGCTCAATGCTCTCCAGGTGACAGTAAACG 3767
Db 8697 TCCATGACACATAGGACATTTCACTGCAAGGTGCCACCGGGTAAATTCCTGCAGG 8756
QY 3768 TCAATATCAG---AGCGAGCATCTGAGAAATTCATGACACCGTGGAGAAAAGATCAG 3824
Db 8757 TCTGATTCAGGACACAGAAACGCGTCCGTCGTCAGAAATACATATCATATGAC 8816
QY 3825 GGAAGTTTGTGAGAGAGAGTACTTGTCCACCCGTCATGAGAAAGCTGTAAAGT 3884
Db 8817 CTCAAACCGGTGGAGAGAAAATTTTCAATTAAGACCACTATGAAAAAGATCCCTT 8876
QY 3885 GCCACGTTTACGACTTGAAGAGACGTGCGCGGTACATTAACATGACAGGCGCAG 3944
Db 8877 GCACACTTATCAACACAGCCACGAGAACCGTGGAGAAATGACATGATATGCGC 8936
QY 3945 GCCACACGCGTATTAAGTCTATCTGAGAGAAAGGTGACGAGGATTAATTAACAC 4004
Db 8937 CAGATAGCCGAGACAGACGTGCTATCAACAGCATCTGGCAATGTAAAGTCA---CAG 8993
QY 4005 CTCTGCGAAGAACGTCACTTACGATGTAAAGTGTGCGACTACAGACAGATCTGTA 4064
Db 8994 TCGGAGAAAAGAGTAAATATACAACTGCACTGTGAAACCGGAAACGTTGGCACTA 9053
QY 4065 GCACGCGAAGAGATGACGCGCTGCACTAAAGCAAAAGTGAATGCTTCAACAGAGC 4124
Db 9054 ATTGGAATGACATCAACGCTG---TCTTAATGAGCAATGTCAGCTTCACTGACG 9110
QY 4125 ACCAAAGAAATGGGTCTTCACTCGCGGATCTTATTAGGACACAGACCACTCACTGC 4184
Db 9111 ACCATTAAGAAATGSCAGTTCACTCCTTTTCCGCCAGAGCCGAGAACCGGCTAGAA 9170
QY 4185 AAGTAATATGACATTCATTCGCTTGAACCCGAGACGTGCGCGGTTCCGTTAAGCTC 4244
Db 9171 AAGGCAAGTTCATATCCATTCCTGTTGCAAACTACATCAATGAGAGTTCAATGCGC 9230
QY 4245 ACACGCTTACAGTACGAAAGTGTTCAAAGGATCACTTCCACTGACTGATGAGAC 4304
Db 9231 GCGAACAACCGTATCAACGCAAAAAGAAAGTACATGCACTTCAACCAATATATC 9290
QY 4305 CAACATTTGTCACACAGAAATTTGGGCTGCGAGACAGCAACAGCAATGATTA 4364
Db 9291 CCAAGCTCTTTTCTTACCGCACATGCGGTGAGACCCGAGATACAGAGAAATGGGTGA 9350
QY 4365 CAGGCTTACATCCAGAAATTTTCTGTGGGCGAGAAAGGCTGAGTACGATGGGCTA 4424
Db 9351 CAGGCGGTGAGACGACCATACCGTACCAATGAGACGGGATGAGATCACTGGGGAA 9410
QY 4425 ACCATGACACAGTACAGATCTGGGCCCCAGAGATGCGACACAGGCGACCATGATGCG 4484
Db 9411 ACAACGACCCAGTGAAGCTTTGTTCACTACACATGAAAGGAAACCGCACGCTGGC 9470
QY 4485 CGCATGAGATCATATCATATTAATCATGCGGATCCAGTCTACATGTATGTGCTGT 4544
Db 9471 CGCATGAGATGATACATATCTATGAGGCTTAAACCGGCGGATGACAGTATCGCGGCTG 9530
QY 4545 GTGTGTGCTCTTGTCTATCTCTGTAGGCACTGCATATACAGACCTTGATGCGCAAAG 4604
Db 9531 TCGGATGAGACTTACTGCGCTTATATCATCTTCGCGGTGTGATCATGCTGTGTCGG 9590
QY 4605 CAGAAAGACTGCGCTGAGCCATPAGCGCTTGACACCGAACGAGTCCACAGCAT 4664
Db 9591 CCGCATTAAGTGTATGACCTTATGCTTAAACACAGAGACTGCACTTCCGTGACGC 9650
QY 4665 TAGCGTTTGTGTGCTGATTCGCGCAACACAGCTGAAACATTTGAGAAATTTGAAC 4724
Db 9651 TGGGATATCTGTGTGCGCCCGCGGCGACAGAGTATGTGAGACATATGCGCT 9710
QY 4725 ATCTGTGTTTAAACAAACCAACCGTTTCTGCGACACAGTTGTGATTCCTTGGCAGCG 4784
Db 9711 ACTGTGGACCAAAACCAAGCGTTGTTCTGTGAGGTTGGCGGCCCTGTTGCTGCA 9770
QY 4785 TTGTTATTCGTTCCGCTG-----TTTCATGCTGATGCGCTTTTAT 4829
Db 9771 TCTCATCATACAGTATTTGCTCAGAAACGCTGCTGCTGTATGAAGCTTTCTTTT 9830
QY 4830 TGGTTGAGCGCTGTGCTGCGGAGAGTACGCGCTTGAACATGCACTGTCGCA 4889
Db 9831 TAGTGTACTAGAGCTGCGGAGCAACCGCAGAGCTTACAAATTCAGATATGCGGA 9890
QY 4890 ATGTTCCGGGATTCGCTATTAAGCGTTGTGCAACGCGAGGTACGCGCCACTTAAC 4949
Db 9891 ACGTGTGGGTTTCCGTTATAGGCTTCACATTAAAGCCAGATATAGCCCTCACTT 9950
QY 4950 TGGAGATCAAGGCTGCTCATTCGAAATTAACACTTCACTAACAGAGATGCTGACT 5009
Db 9951 TGCAGATGACAGTGTGTTGAAACACAGCTCGAACCAACCTTAAATTTGAAATACATACT 10010
QY 5010 GCAAAATTCACACAGTCAATCTTCAACCAAGTTAAATGCTGCGGCTCCCTGAGTGA 5069
Db 10011 GTGAGTACAAAGAGGTGTCCCGTCCGTAAGTGAAGTCTGCGGCGCTCAGAGTGT 10070
QY 5070 AAGCATCTTAAAGCGGATTAACATGCGCGCTTTTGGCGGTGTACCTTTATGT 5129
Db 10071 CCACTAAAGAAAGCTGACTACCAATGCAAGGTTTACACAGGCGGTGTACCTTATGT 10130
QY 5130 GGGGAGGCGACATGCTTCTGAGACGTGAGAAACACAACTGAGTGAAGCGGTAGCTG 5189
Db 10131 GGGAGGGGCAATTTGCTTCTGAGCTCAGAAAACAGCACTCAGGAGGCTTACGTG 10190
QY 5190 AGTTGCTCCAGACTGCACTATATGATCAACGAGTGCACCTTAAAGTTACACAGCTGCTC 5249
Db 10191 ATGATGACAGATGATGAGGATGATCAGCATCTGTTTCAAAAGCCCATACAGATGCG 10250
QY 5250 TGAAGTCCGCTGCGTATGATATGAGCAACACACCGGACACTGAGATAGCTTTGTCA 5309
Db 10251 TGAAGGCAAGTGAAGGTTATGTACGCAACCTTAAACAGAGTGTGAGATTTACCTGA 10310
QY 5310 ATGGCGTCAAGCAAGGTTCTCAACGAGACTGAAAGTCATAGCAGGCGGATATACGCG 5369
Db 10311 ACGGAGACATGCGCTGACATAGGGGATCTCAATTTATTTGCGGCGCTGTATCGG 10370
QY 5370 CTTTTCACCTTTGACATTAAGTGTCTATCAAGAAAGGGCTTGTTTACATACGACT 5429
Db 10371 CTTGACCCGCTTGAACAAATAGTGTGTATCAAAAGCAAGTGTCAATACGAGACT 10430
QY 5430 TCCCTGAGTATGAGCTATGAACACAGAGCGTTTGGCGATATTCAGCATCTCGCTTG 5489
Db 10431 TCCGCGCTTACGATCTGGGCAACAGGGCGTCTCGCGCATTCAAAGCAAGCATGTG 10490
```

QY 5490 ATGCTACAGACATAGTACCGCCGACTGACATACGCGCTGCTGAGACCTTCTGTCAAGACA 5549
DB 10491 AGAGTAAGACACTGTACCGGAACACGGCACTGAGCTGGACAGCCCTTACCCGGCATGG 10550
QY 5550 TCCAGTCCCTTACACCCCAAGAGATATGAGGTATGAATGTGAAAGAACTCAAGAC 5609
DB 10551 TCCAGTACCGGTACACAGACGACCTTACGGGTTCAMATATGGTTAAAGGAAAAAGGA 10610
QY 5610 GACCCCTGCAAGAAACAGACACCATTTGATGTAAATTTGAAGTGGAGCCCTGCGAGCGT 5669
DB 10611 CAGCCCTAAATACGAAAGCTCTTTTGGCTGCCAAATCAAAAGAACTTTCAGAGGCA 10670
QY 5670 CTAACTGTGCTTACGGGACATCCCTATCTGATTTGACATCCCTGATGAGCTTTTGTGA 5729
DB 10671 TGAATGCGCGGTGGAAACATCCCTGTCTCATGATTTGCTGACAGCGCTTTTACCC 10730
QY 5730 GATCATCAGATACCAACATTTTGAAGTTAGCTGCAAGTACAGACCTGCAATTTAT 5789
DB 10731 GCATTTGAGGCGCGGACCATATTGACCTGACCTGACAGTGGCTACCTGACGCACT 10790
QY 5790 CTGACACTTTGGTGTCTCTAATTAACAGTACAAAGCTGACAGGAGGAGCATTTGTC 5849
DB 10791 CTTGGATTTTGGCGCGCTCTTACACTGACGTAACAAAGCAAGAGGAGGAGCTCT 10850
QY 5850 CAGTTCACTCCCACTCCAGACAGCTGTTTGAAGAAAGCAACACATGTAAGTGGCG 5909
DB 10851 CTGTACACTCGCACTTAACGTAAGTACTCTACAGAGGCGCACAGAAAGTGAAGACAG 10910
QY 5910 TAGGCACATTAACACTATTATTAACATGAGCCCAAGCAAAATTTATAGTTTGGC 5969
DB 10911 CAGGTAGGTGACCTTAACACTCTCCAGGCGAAGGACATCACTTTTGGTGGTGGC 10970
QY 5970 TATGGGGAAGAGTCACTGACATGCTGAATGTAAACCCGGGCGACCAACATATTG 6029
DB 10971 TATGAGAGTCTAGGCGCACCTGTTAGCGTGTGAGGCCCCGAAAGCAACATAGTCC 11030
QY 6030 GAGAACCATTAATTCAGACCAAGATTCAGAGCGGAGTTTCCAAAATCTTGAAGT 6089
DB 11031 CATATGCGGCTAGCACAGTAAGTATGTTTCCAGACATGTCCGGGACCCGACATATCAT 11090
QY 6090 GCGTGTCTTCACTGTTTGGGGAGCATCATCCCTCATTTGTGTAGACTTATAGTTGG 6149
DB 11091 GGGTGCAGAAATCTCGGGTGTCTGGGGGCTTTCGCAATCGGCGCTATCCGTGGTGG 11150
QY 6150 T 6150
DB 11151 T 11151

RESULT 6
US-09-991-258-1
; Sequence 1, Application US/09991258
; Patent No. US20020141975A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Robert
; APPLICANT: Keith, Paula
; APPLICANT: Dryga, Sergey
; APPLICANT: Caley, Ian
; APPLICANT: Maughan, Maureen
; APPLICANT: Johnston, Robert
; APPLICANT: Davis, Nancy
; APPLICANT: Swanson, Ronald
; TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOSES WITH MODIFIED HIV GENES FOR USE
; TITLE OF INVENTION: VACCINES
; FILE REFERENCE: 01113.0001U3
; CURRENT APPLICATION NUMBER: US/09/991.258
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 09/902.537
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216.995
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1
; LENGTH: 12523
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. US20020141975A1e =
US-09-991-258-1
Query Match 14.3%; Score 926.2; DB 10; Length 12523;
Best Local Similarity 68.1%; Pred. No. 2.2e-265;
Matches 120; Conservative 0; Mismatches 613; Indels 6; Gaps 2;
QY 514 AACACACAGAGTGCCTAGAGCATCTGAGAGGCGAGTGGAGAGTCAATCCGTCAC 573
DB 5617 ACCCGCAGCGGTAAATGGGTATTTACAGAGAGAGATTTGAGCGCTTGTAGCACAC 5676
QY 574 ACTCCAACTGACGGTATGACCGGAGCGTATATTTCTCATTCGAAACAGGCCAAGTTC 633
DB 5677 A--ACATGACGGTTTGTATGGGTGATCATCTTTTCTCCGACACCGGTCAAGGCG 5733
QY 634 ACCTTCAACAGAAATCACTAGTCAATGTAACTTCAAGAACTTATTTGATCGGGCG 693
DB 5734 ATTTACAAACAAATCAAGTAAGCAACGCTGTATCCGAGTGGTGTGAGAGAGACCG 5793
QY 694 TCCATGAGAGTATTAAGCGCGCGCGCTCGATCTCGAAGAGAGAAATGTTACAGAGA 753
DB 5794 AATTGAGATTTTCGTATGCCCGCGCTCGACAGAAAGAAAGAAATTTACTACGCAAG 5853
QY 754 AACTGCAATTTAGCGCTCTGAAAGAAATAGAGAGGATTCATACAGAAAGTAGAAA 813
DB 5854 AATTAGAGTTAAATCCCACTGCTTAACAGAGAGATACAGTCCAGAAAGTGGAGA 5913
QY 814 ATATGAAGCAATTAACAGGAGGAGGAGTCTATTCTGATGGGCAATATCTATCATGAG 873
DB 5914 ACATGAAGGCAATTAACAGTACGATATTCGCAAGGCTTAGGGCATATTTGAAGGAG 5973
QY 874 AAGTGAATCCGTGAGGTGATACAGAGTCAATTTCCGTGCAATCTACCTGCTCAAG 933
DB 5974 AAGGAA--AGTGAAGTCTACCGAACCTGCAATCTGTTCTTGTATTCATCTAG 6030
QY 934 TAAATTAAGGTTTACATCTGTCAGAGGTGCGGTTAAACGTGCAACTTATTTCAAG 993
DB 6031 TGAACCGTCTTTTCAAGCCCAAGGTGAGAGTGAAGCTGTAAACCCATGTTGAAG 6090
QY 994 AGAATTAACCTTACAGTACGACGATTTATGTATTAACAGATGAATACGATCGTATCTTGA 1053
DB 6091 AGAATTTCCGAGTGGCTTCTTACTGTATTTATCCAGAGTACGATGCTATTTGAGCA 6150
QY 1054 TGGTGAAGCGGCAATCGTGTCTGTATTAATACGACCACTTTTGTCCGGCTAACTGAGAA 1113
DB 6151 TGGTTGAGGAGCTTCAAGTGTCTTGAACATGTCAGATTTTGGCTTGCAGAAAGTGGCCCA 6210
QY 1114 GCTACCCCAAGACATATGATTTTGGAGCCAGAGATTAAGATACGCGCTCCATCGCTTA 1173
DB 6211 GCTTTCCAAAGAAACATCTTATTTTGAACCAACATTAAGATCGGAGTGCCTTCAAGCA 6270
QY 1174 TACAGATTAATTAACAAATGTATTTGCTGAGCTACTAAAGAAATTTCAACGTTACCC 1233
DB 6271 TCCAGAAACAGCTCCAGAACGTCCTGGAGCTGACCAAAAAGAAATTTGCAATGTCAGC 6330
QY 1234 AAAAGCAGAAATTAACGTTCTTAAATTCGGGCGCATTTAATGTTGATTTTCAAGAAAT 1293
DB 6331 AAAAGAGAAATTTCCGTATTTGATTTGAGGCGCTTTAAATGTGAAGTCTTCAAGAAAT 6390
QY 1294 ACGATCAATGATGAGTACGATGATACCTTTCGCAATACCTATTCGGCTAATCAAG 1353
DB 6391 ATGCGTGTAAATTAATGATTTTGGAAACGTTTAAAGAAACCCATCAGGCTTACTGAAG 6450
QY 1354 AGAAGTTAGCAATATGTGCAAAAGCTGAAAGGCGGAAAGCAGACGATTTGTTGGA 1413
DB 6451 AAAACGTGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6510

1414 ATACTATAATCTAAACCGTTGACGAGATATACCAATGATCAATTCGTATGATCTAA 1473
6511 AGACACATAATTTGATATATGTTGACGAGCATACCAATGAGACAGTTTGTATGACCTTAA 6570
1474 AGAGAGATGTCAAAGTTACTCCCGCACGAAACATACAGAGAGCGGCTTAAGGTGACG 1533
6571 AGAGAGACGTGAAGATCACTCCAGGAAACAAAACATCTGAAAGAACGCGCCCAAGGTACAG 6630
1534 TTATTACGGCTGACGATCCCGCTTCCCGCTTACCTTTGCGGGATCCATCGGGAATTAG 1593
6631 TGATCCAGGCTCGCATCCGTCAGCAACAGCGATCTGTGCGAAATCCACCGAAGCTCG 6690
1594 TCCGTAGACTGAATGCGGTGCTTGCCTTCAATATCACTCTCTGACATGTACGCG 1653
6691 TTGAGAGATTAAAGCGGCTCTGCTTCCGAACATTCACTGTTGATATGTGCGGTG 6750
1654 AAGATTTTATGATGATTTATGCTGAACATTTCCACACGCGGACCCAGTATGGAACGG 1713
6751 AAGACTTTGACGCTATTATAGCCGAGACCTTCAGCCGCGGAGATTGTCTGGAACCTG 6810
1714 ACATCGGCTGTTGATTAAGGGAAGACGAGCGCTATCGCATTTCCGCGCTTGTGATCC 1773
6811 ACATCGGCTGTTGATTAAGGGAAGACGAGCGCTATCGCATTTCAACGCGTTAATGATTC 6870
1774 TTGAGACTTATGATGTGACCAACCGCTCTTATGATTAAGAGCGCGCTTCCGCAATA 1833
6871 TGGAGACTTATGATGTGACCAACCGCTCTTATGATTAAGAGCGCGCTTCCGCAATA 6930
1834 TCACATCTGTGACCTTACCTACAGAAACGAGCTTAAATTTGTTGCTCATGTGAATCCG 1893
6931 TTTTCATCAATCAATTTGCTCCCACTAAACTAAATTTAAATTCGAGCCATGTGAATCTG 6990
1894 GTATGTTCTTAAGCGCTGTTGTCAACACACTAGTCAATATCATATGTTGTCAGAGTAC 1953
6991 GAATGTTCTTCAACCTGTTGTGAACACATGATTTACATTTGATTCAGAGAGTGT 7050
1954 TACGTGAACGCTTAAACGACGCTGACGCGCTGCGGCTTATGCGGACGATTAACATGTC 2013
7051 TGAAGAAACGCTTAAACGCGGATCAACATGTGACGATTCATTGAGATGACAAATTCGGA 7110
2014 ATGCTGTGCTTCCGACACCTTGAATGCGGAGATGCGCACTTGGCTGAACATGGAAG 2073
7111 AAGAGTCAATTCGACAAATTAATGCGACAGGTGCGCACTGTTGAATATGGAAG 7170
2074 TAAATAATATGATGAGTATGTTGATTAATCAAGACCTTACTGTTGCGGGGATTTATCC 2133
7171 TCAAGATTATGATGCTGTGTGCGGAGAAAGCGCTTATTTCTGTGAGAGGTTTATTT 7230
2134 TGTGTGACACGATTAACAGGACAGCTGACAGTGCAGACCTCTTAAAGGCTTTTAA 2193
7231 TGTGTGACTCGGTGACGCGGACAGAGGTGCGGTGCGGAGACCCCTTAAAGGCTTTTAA 7290
2194 AGCTTTGAAAACATTTGCGAGTCAATGATACCAAGACTGCGACCGCGCGGCACTGC 2253
7291 AGCTTTGAAAACCTTGGCAGACAGATGATGATGATGATGATGATGATGATGATGATG 7350
2254 ATGATGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2313
7351 ATGATGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7410
2314 AATCCAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2373
7411 AATCAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7470
2374 AAGAGCTTAAAGATTCAGAGATTAAGAGGAGCCCAATCACTCTTACGCGTGAAGCTA 2433
7471 GCAATGTTAAATCAATTCAGATTAAGAGGAGGCGCCCTTAACTCTTACGCGTGAAGCT 7530
2434 AATAGTGAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2493
7531 AATGATGAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2550

Db 7531 AATGATGAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG

RESULT 7
US-09-991-258-14
Sequence 14: Application US/09991258
Patent No. US20020141975A1
GENERAL INFORMATION:
APPLICANT: Olmsted, Robert
APPLICANT: Keith, Paula
APPLICANT: Dryga, Sergey
APPLICANT: Maughan, Ian
APPLICANT: Johnstone, Robert
APPLICANT: Davis, Nancy
APPLICANT: Swanson, Ronald
TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOSES WITH MODIFIED HIV GENES FOR USE
FILE REFERENCE: 01113.000103
CURRENT APPLICATION NUMBER: US/09/991.258
CURRENT FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: 09/902,537
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: 60/216,995
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FASTA for Windows Version 4.0
SEQ ID NO 14
LENGTH: 12379
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE: Description of Artificial Sequence; No. US20020141975A1 =
OTHER INFORMATION: synthetic construct
US-09-991-258-14

Query Match 14.3%; Score 924.6; DB 10; Length 12379;
Best Local Similarity 68.0%; Pred. No. 6.5e-265;
Matches 1319; Conservative 0; Mismatches 614; Indels 6; Gaps 2;
514 AACACGAGAGTGTGACGATCTCAGAGGAGGAGTTGACGATGATCCGTCAAC 573
5617 ACCCGCAGCGTAAATAGGATGATTAACAGAGAGGAGTTGAGCGCTTGTGACACAC 5676
574 ACTCACTAGCGTATGAAGCGGAGCGTATATTTCTCATGCGAAACAGGCGAAGTGC 633
5677 A--ACAATAGCGTGTGATGCGGCTGATATCAATCTTCTCCGACACGCGTCAAGGCG 5733
634 ACCTTCAAGCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 693
5734 ATTACAAACAAATTCAGTAAAGGCAACGCTGATCCGAAAGTGTGTGAGAGGACCG 5793
694 TCCATGAGAGTATTAAGCGCGCGCTGATCTCGAAGAGAGAAATGTTACAGAGA 753
5794 AATTGAGATTTCTGATGCGCGCGCTGATCCGACCAAGAAAGAAATTAATCTACGAGA 5853
754 AACTGAATTAATGCGCTTGAAGAAATGAAGACAGTATCAATCAAGAAAGTGAAGA 813
5854 AATTACAGTAAATCCACACCTGCTAAGAGCAATGATCCAGAGAGGTGAGA 5913
814 ATATGAAGCAATTAACGCGAGGACTATTTCTGATGATGATGATGATGATGATGATGATG 873
5914 ACATGAAGGCAATTAACGCTGATGATGATGATGATGATGATGATGATGATGATGATG 5973
874 AATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 933
5974 AAGGAAA--AGTGAATGCTACGAGACCTGCAATCTCTTGTGATTTCACTGATG 6030
934 TAAATTAACAGTTTATCACTGACAGAGTGGCGTTAAACGTGCACTTATGATGATGATG 993
6031 TGAACGCTGCTTTTCAAGCCCAAGGTCGACAGTGAAGGCTGTAAGCCATGTTGAAG 6090
994 AGAATTTACCTACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1053
6091 AGAATTTACCTACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6150

QY 1054 TGTGGACGGCGCATGCTGCTGTAGATACAGCCATTTTGTCCGGCTAAACTGAGAA 1113
 DB 6151 TGTGTGACGGAGCTTCATGCTCTTATGACACTGCGAGTTTGGCCCTGAAAGCTGCGGA 6210
 QY 1114 GCTACCCAAAGAGCATAGCTATTGTCAGCCAGAGATTAAGATCAGCCGTCCATGCGCTTA 1173
 DB 6211 GCTTCCAAAGAAACACTCTATTGTTGAAACCCACATATCGATCGGAGCTTTCAGGGA 6270
 QY 1174 TACAGATATCATTTACAAATATGTTGGCTGACGCTACTATAAAGAAATGCAACGTTACCC 1233
 DB 6271 TCCAGAACACGCTCCAGAAAGCTCTGGAGCTGCCACAAAGAAATGCAATGTCACGC 6330
 QY 1234 AAATCGAAGATTACCTGTCTTATGATTCGCGCGCATTTAATGTTGATTGTTCAAGAAAT 1293
 DB 6331 AAATGAGAAATGTCGCCGATTTGATTTGGCGGCTTTAATGTAAGATGTTCAAGAAAT 6390
 QY 1294 ACGCATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1353
 DB 6391 ATGCGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6450
 QY 1354 AGAAGCTTACGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1413
 DB 6451 AAAAGCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6510
 QY 1414 ATACTCATTAATCTAAACCGTTGACAGATACCAATGATCATTTGTCATGATCTAA 1473
 DB 6511 AGACATTAATTTGAATATGTTGACAGACATACCAATGACAGTTTGTAAATGACTTAA 6570
 QY 1474 AGAGAGATGTCAAAGTTACTCCCGCAGAAACATACAGAGAGCGGCTTAAGTGCAG 1533
 DB 6571 AGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 6630
 QY 1534 TTATTCAGGCTGACATCCCGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1593
 DB 6631 TGATTCAGGCTGACATCCCGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 6690
 QY 1594 TCCGTAAGTGAATGCGGTGCTTTCGCAAAATATCATCTCTTTCACATGATGACGG 1653
 DB 6691 TTAGAGATTAATTCGGCTCTGCTTCGCAAAATATCATCTCTTTCACATGATGACGG 6750
 QY 1654 AAGATTTGATGCGATTAATGCTGAACATTTTCACACGCGGCGCAGTATGAAACGG 1713
 DB 6751 AAGCTTTGACGCTTAATGCTGAACATTTTCACACGCGGCGCAGTATGAAACGG 6810
 QY 1714 ACATCGCGCTGTTGATTAAGCGGAAGACGATGATGATGATGATGATGATGATGATGATGAT 1773
 DB 6811 ACATCGCGCTGTTGATTAAGCGGAAGACGATGATGATGATGATGATGATGATGATGATGAT 6870
 QY 1774 TTGAGGACTTAGGCTGACCAACCGCTCTTATGTTGATGAGGCGGCTTCGCAATA 1833
 DB 6871 TGGAGAGCTTAGGCTGACCAACCGCTCTTATGTTGATGAGGCGGCTTCGCAATA 6930
 QY 1834 TCAATCTGTGACCTTACCAAGAACGAGTTTAAATTTGTCGCAATGATGAATCG 1893
 DB 6931 TTTTCATCAATCATTTGCCCACTAAACTTAAATTTAAATTTGCGACCATGATGAATCG 6990
 QY 1894 GTATCTTTTAAACGCTTTTTCGCAACACTGATGATGATGATGATGATGATGATGATGATGAT 1953
 DB 6991 GAATCTTTCTCACTGTTTTCGCAACACTGATGATGATGATGATGATGATGATGATGATGAT 7050
 QY 1954 TACGTAACGCTTAAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2013
 DB 7051 TGAAGAACGCTTAAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7110
 QY 2014 ATGCTGCTGCTCCGACACTTGTATGCGGAGATGCGCACTTGCTGAACATGAG 2073
 DB 7111 AAGGAGTCAAAATCGCAAAATTAATGCGACAGATGCGCACTTGCTGAACATGAG 7170
 QY 2074 TAAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2133
 DB 7171 TCAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7230
 QY 2134 TGTGTGACCAAGTAAACAGGACAGCTGCGAGAGTGCAGAACCTCTAAAGAGCTTTTAA 2193

DB 7231 TGTGTGACTCCGTGACCGGACAGCGTCCGCTGTGCGAGACCCCTTAAAGAGCTGTTTA 7290
 QY 2194 AGCTTGGAAACCATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2253
 DB 7291 AGCTTGGAAACCATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7350
 QY 2254 ATGATGAACATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2313
 DB 7351 ATGATGAACATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7410
 QY 2314 AATCGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2373
 DB 7411 AATCGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7470
 QY 2374 AAGGCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2433
 DB 7471 GCAAGTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7530
 QY 2434 AATAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2493
 DB 7531 AATAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7590

RESULT 8

US-09-991-258-17

Sequence 17, Application US/09991258

Patent No. US20020141975A1

GENERAL INFORMATION:

APPLICANT: Olmsted, Robert

APPLICANT: Keith, Paula

APPLICANT: Dryga, Sergey

APPLICANT: Caley, Ian

APPLICANT: Maughan, Maureen

APPLICANT: Johnston, Robert

APPLICANT: Davis, Nancy

APPLICANT: Swanson, Ronald

TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOMES WITH MODIFIED HIV GENES FOR USE

TITLE OF INVENTION: VACCINES

FILE REFERENCE: 01113.0001U3

CURRENT APPLICATION NUMBER: US/09/991,258

CURRENT FILING DATE: 2001-11-16

PRIOR APPLICATION NUMBER: 09/902,537

PRIOR FILING DATE: 2000-07-09

PRIOR APPLICATION NUMBER: 60/216,995

NUMBER OF SEQ ID NOS: 19

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO. 17

LENGTH: 13584

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence; No. US20020141975A1e =

US-09-991-258-17

Query Match 14.3%; Score 924.6; DB 10; Length 13584;

Best Local Similarity 68.0%; Pred. No. 7e-265;

Matches 1319; Conservative 0; Mismatches 614; Indels 6; Gaps 2;

QY 514 AACCAACGAGTGGCTGATGACGATCTGAGAACGCGAGATTGAGACGATGATCCGTCAAC 573
 DB 5617 ACCGCGACGGCTGAATGAGGTGATTAACAGAGAGAGATTGAGCGGTTGATGACGACAC 5676
 QY 574 ACTCAACTGACGATGATGAGCGGAGCGTATTTTCTCATCGGAACAGGCCAAGGTC 633
 DB 5677 A---ACATGACGCTTGTGATGCGGGTGCATACATCTTTCTCCGACACCGGTCAAGGCG 5733
 QY 634 ACCTTCAACAGAAATGAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 693
 DB 5734 ATTTCACAAACAAATTAAGTAAAGCAACGCGTATTCGAAAGTGTGATGAGAGACCG 5793

694 TCATGAGAGATATTAGCCCCCGCGCTCGATCTCGAAAGAGAGAAAATTGACAGAGA 753
5794 AATTGGAGATTTTCGTATGCCCCCGCGCTCGACCAAGAAAAGAAATTAAGACAGA 5853
754 AACTGCAATTATGCGCTCTGMAAGAAATAGACAGGATCAATCAAGAAAGTAGAAA 813
5854 AATTACAGTTAAATCCACACCTCTTAACAGAAACAGATACAGTCCAGGAAGGTGAGA 5913
814 ATATGAAGCAATTACAGCGGAGGAGCACTATTCTGATGGGACATATCTATCAACAG 873
5914 ACATGAAGGCATTAACAGCTAGACGTATCTGCAAGGCTTAGGGCAATTAATTGAAGCAG 5973
874 AAGTAATCTGTGAGGTGTATACAGAGTCAATTAATCTGTACATCTGCAACG 933
5974 AAGGAAA--AGTGAAGTGTACCGAACCCTGCACTCTGTTCTTTGATTTCACTAGTG 6030
934 TAATTAACAGGTTTACATCTGCAAGGTGCGGTTAAACGTGCACTTAAGTTACAG 993
6031 TGAACCGTCTTTCAAGCCCCCAAGGTGCGAGTGAAGCCTGTAAAGCCATGTTGAAG 6090
994 AGAATTACCTAGAGTACAGTACGATATGTATTAACAGATGAATAGATGCGATCTTACA 1053
6091 AGAATTTCCGAGCTGCGCTTCTTACTGTATTAATTCAGAGTACGATGCTATTTGACA 6150
1054 TGGTGAAGCGGCGCATCGTGTCTGTATAGATACAGCCACTTTTGTCCGGCTAACTGAAA 1113
6151 TGGTGAAGCGAGCTTCACTGTCTGTATAGACTGCAAGTTTGTCCCTGCAAGAGTGGCA 6210
1114 GCTAACCAAGAGCATAGCTATTGTGCAAGCAGATGAATGATCAAGCCGTCCATCGCTTA 1173
6211 GCTTTCCAAAGAAACACTCTTATTTGAAACCCCAATACGATCCGAGTGTCTTCAACGA 6270
1174 TAGAGATATCTTCAAAATGTAATTTGCTGAGCTACTTAAAGAAATTTGCAACCTTACC 1233
6271 TCCAAACACGCTCCAGAACGCTCTGAGAGTGCACAAAAGAAATTCATATCTACGC 6330
1234 AATGCGAGAAATTAATCTGTCTTATGATTCGCGGCAATTAATGTTGATTTTCAAGAAAT 1293
6331 AATGAGAGAAATTTCCGCTATTTGATTTGCGGCGCTTAAATGTTGATTTTCAAGAAAT 6390
1294 ACGCATCAATGATGAGTACTGGGATACCTTTGCGCATACCTTATTCGGCTAACTACAG 1353
6391 ATGGGTAAATTAAGATATTTGGGAAACGTTTAAAGAAACCCCATCAGGCTTACTAGAG 6450
1354 AGAAGCTTACGCAATATGTCAGAACGTCGAAGGCGGAAAGCAGCAGCATTTGTCGA 1413
6451 AAAACGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6510
1414 ATACTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1473
6511 AGACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6570
1474 AGAGAGATGTCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1533
6571 AGAGAGATGTCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6630
1534 TTATTCAGGCTGAGATCCCTTGTACCCGCTTACCTTTGCGGATTCATTCGGAATTAG 1593
6631 TGATTCAGGCTGAGATCCCTTGTACCCGCTTACCTTTGCGGATTCATTCGGAATTAG 6690
1594 TCCCTAGACTGAATGCGGCTTGTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1653
6691 TTAGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6750
1654 AAGATTTGATGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1713
6751 AAGATTTGATGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6810
1714 ACATGCGGTGCTTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1773
6811 ACATGCGGTGCTTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6870
1774 TTGAGACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1833

6871 TGGAAGCTTAAGGTGAGACGAGAGCTGTGACGCTGATTTGAGGCGGCTTTGCGCGAAA 6930
1834 TCACATCTGTCACCTTACCTACAGAAACGAGGTTTAAATTTGGTCCATGATGAAATCCG 1893
6931 TTTCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6990
1894 GTATGTTCTTAACGCTGTTTGTCAACACTAGTCAATATCATGATTTGCTAGCAGACTAC 1953
6991 GAATGTTCTCAACCTGTTTGTCAACACTAGTCAATATTAATTTGATGCAAGCAGAGTGT 7050
1954 TACGTAAACGTTAAACCACTGACAGCTGCGGCTCTATTCGCGAGATPAACATAGTGC 2013
7051 TGAGAGAACGCTTAAACCGATACACATGTCACATTCATTTGAGATGACAAATATCGTGA 7110
2014 ATGCTGCTGCTCCGACACCTTATGCGGAGAGATGCGGCACTGCTGCTGAACATGGAAG 2073
7111 AAGAGTCAATGCGAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7170
2074 TAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2133
7171 TCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7230
2134 TGGTGAACGAGATPAACAGGACAGCTGACAGTTCGACACCTCTTAAAGGCTTTTAA 2193
7231 TGTGTACTCCGTGACCGGACAGCGTCCGTGTGGACAGACCCCTTAAAGGCTGTTTA 7290
2194 AGCTTGAAGAAACATTTGCCAGTGTATATCCAAAGCTGCAAGCCGCGCGGCACTGC 2253
7291 AGCTTGGCAAACTCTGCGAGCAGAGCATGATGATGATGATGATGATGATGATGATGATGATG 7350
2254 ATGATGAAGCAATGCGATGGAACAGAAATTTGAATTAACGAGATTAAGTGAAGGCTGAG 2313
7351 ATGAGAGTCAACAGCTGGAACCGATGAGTATTTCTTGAAGCTGTGAAGGCTGAG 7410
2314 AATTCAGATPAACGATCATATCTGCGAGGCTGTATCATACGCTTCTGTCCAGCTTAAAG 2373
7411 AATCAAGATTAAGAAACCGTGAAGAACTTCATCATAGTATTAATTAATTAATTAATTAATTA 7470
2374 AAGCGTTAAGAACTTCAAGATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2433
7471 GCAAGTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7530
2434 AATAGTGAAGTATGAG 2452
7531 AATGACTACGACATAGTC 7549

RESULT 9
US-09-991-258-2
Sequence 2, Application US/09991258
Patent No. US20020141975A1
GENERAL INFORMATION:
APPLICANT: Olmsted, Robert
APPLICANT: Keith, Paula
APPLICANT: Dryga, Sergey
APPLICANT: Cauley, Ian
APPLICANT: Maughan, Maureen
APPLICANT: Johnston, Robert
APPLICANT: Davis, Nancy
APPLICANT: Swanson, Ronald
TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOSES WITH MODIFIED HIV GENES FOR USE
TITLE OF INVENTION: VACCINES
FILE REFERENCE: 01113.000103
CURRENT APPLICATION NUMBER: US/09/991.258
CURRENT FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: 09/902,537
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: 60/216,995
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2

Publication No. US20030096397A1
GENERAL INFORMATION:
APPLICANT: Dubensky Jr., Thomas W.
Folio, John M.
Bell, Barbara A.
Schlesinger, Sondra
Dryga, Sergey A.
Frolov, Ilva
TITLE OF INVENTION: RECOMBINANT ALPHAVIRUS-BASED VECTORS
WITH REDUCED INHIBITION OF CELLULAR MACRO-MOLECULAR
SYNTHESIS
NUMBER OF SEQUENCES: 125
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed Intellectual Property Law Group PLLC
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/507,362
FILING DATE: 18-Feb-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 930049.457D6 /1196.011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
LENGTH: 8000 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 102:
US-09-507-362-102
Query Match 12.8%; Score 827.6; DB 11; Length 8000;
Best Local Similarity 61.9%; Pred. No. 5.9e-236;
Matches 1390; Conservative 0; Mismatches 829; Indels 25; Gaps 4;
QY 580 ACTGACGGTATGAAGCGGAGCGTATATTTCTCATCGGAAACAGGCCAAGTCACCTTC 639
DB 5746 ACTACTACCGGGGTAGTGGTACATATTTTCAGCGGACACAGGCCCTTGGGACATTGC 5805
QY 640 AACGAATCAGTACGTCAATGTAACTACAAGAACCTATATTGGATCGGCGCTGCATG 699
DB 5806 AAAAGAGTCGTTCTGCAGAACAGCTTACAGAACGACCTTGAAGGCAATGCTCTGG 5865
QY 700 AGAAGTATAGCCCCCGCTCGATCTCGAAAGAGAAATGTATACAGAAAGAACTGC 759
DB 5866 AAAAATTCATGCCCGGTGCTGCACAGTGAAGAAAGAACTCAACTCAGGTACC 5925
QY 760 AATATGCGCTCTGAAGAAATGAAGCAGTATCAATCAGAAAGTAGAAATATGA 819
DB 5926 AGATGATGCCCAACCGAAACCAAAAGTAGTACCACTCTGTAAAGTAGAAATATCA 5985
QY 820 AAGCAATTCAGCGGAGGAGCTATTTCTGATTTGGGCACTATCTATCATCAGAGTGA 879
DB 5986 AAGCATTAACCACTGAGGAGTACTGTACAGACTACGACTGTATTAAGTCTCCACAGATC 6045
QY 880 ATCTGTGAGTGTATACAGATCAATATCTGTACCAATATCTATCGCAACGGTAATTA 939
DB 6046 AGCCA--GAATGTATTAAGATCACTATCGAAACCAATTGTACTCCAGTGGCTACCGG 6102
QY 940 ACAGGTTTACATCTGCAGAGGTGCGGTTAAAGTGCACCTTATGTTATCCAGAGATT 999

DB 6103 CGAACTACTCGATCCACAGCTGCGCTAGCTGTCTGTAACAATCACTACTGCATGAGAACT 6162
QY 1000 ACCCTACAGTACCGCAGTATTGTATTAACAGATGAATTCATGCGTATCTGCATGCTGC 1059
DB 6163 ATCCGACAGTACGATCTTATCAGATTACTACAGGATACGATGCTACTTGGATTTGGTAG 6222
QY 1060 ACCGGCATGCTGTCTGATTAACAGCACTTTTGTCCGGGCTAAACTGGAAGTACC 1119
DB 6223 ACCGGACATGCTGCTGCTGATTAACAGCACTTTTGTCCGGGCTAAACTGGAAGTACC 6282
QY 1120 CAAGAAGCAGTACCTATTTTCAGCGAGATTAAGATCAGCCGCTCCATCGCTATACAGA 1179
DB 6283 CGAAAAAATGAGTATAGAGCCCGGAATATCCGAGATGGGTTCCATCAGCATGACAGA 6342
QY 1180 ATACATTACAATAATGATTTGGCTGACGCTACTAAAGAAATTTGCAAGCTTAACCAATGC 1239
DB 6343 ACACGCTACAAAATGTCTCATTTGCCCACTAAAGAAATTTGCAACGTCACGACGATGC 6402
QY 1240 GAGAAATTAACCTGCTTAGATTGCGGCGCATTTATGTTGATGTTTCAAGAAATAGCAT 1299
DB 6403 GTGAATCGCAACACTGAGCTCAGCGACATTCATGATGCAATGCTTTGCAAAATATGCAT 6462
QY 1300 GCATGATGAGTACTGCGATACCTTTTCGCGATTAACCTATTCGCGTAACTACAGAGAACG 1359
DB 6463 GTATATGCGAGTATTGGGAGAGATGCTGCGAAGCCAAATTAGGATTAACACTGAGTTTG 6522
QY 1360 TTACGCAATATGTCACAAAGCTGAAGAGGCGGAAAGCAGACGATTTGTTCCGAAATCTC 1419
DB 6523 TCACCGCATATGATGCTAGACTGAAGAGGCGCTTAAGGCGCGGCACTAATTTGCAAAAGCGT 6582
QY 1420 ATATCTAAACCGTTGCGAGAGTACCAATGATCAATTCGTCATGATGATCTAAAGAGAG 1479
DB 6583 ATATTTGCTCCATTTGCAAGAAAGTGTCTTATGATGATTTGTCATGATGATGATGATGATG 6642
QY 1480 ATGTCAAGTATCTCCCGGACAGAAACATACAGAGAGCGGCTTAAGTGCAGGTTATTC 1539
DB 6643 ACGTGAAGTTACACGAGGACAGAAACACAGAGAAAGAACCGAAAGTCAAGTGTATAC 6702
QY 1540 AGGCTGCAATCCCTTTGCTACCGGCTTACCTTTGCGGAGATTCATCGGGAATTTAGTCCGTA 1599
DB 6703 AAGCCGAGAAACCCCTGGCCACATGCTTACTTAAGCGGAGTTTACCGGGAATTTAGTGGTA 6762
QY 1600 GACTGAATGCGGCTCTGCGCAAAATATCATCTCTTTCGATGCTGAGTGCAGGGAATTT 1659
DB 6763 GGCCTTACGCGGCTTCTGCTTCCAAACATTACACGCTTTTTCGATGCTGCGGGAAGATT 6822
QY 1660 TTGATGCGATTAATTTGTAACATTTTCCACACGCGGCAACCGATTTGGAACGACATCG 1719
DB 6823 TTGATGCAATCATAGAGAAACATTTCAAGCAAGGCGACCGGTAAGTGAACGGAATTCG 6882
QY 1720 GGTGCTTTGTAAGAGGAAGCAGACGCTATTCGCAATTTGGCGGTTGATATCTTTGAGG 1779
DB 6883 CATCATTTGCAAAAGCAGAGCAGCGCTATGCGTTAAACCGGCTGATATCTTTGAGG 6942
QY 1780 ACTTAGGTGCAACCAACGCGCTCTGATTTGATGAGGCGGCTTCGCAATATCAT 1839
DB 6943 ACCGSGGTGATCAACCACTACTGACTTGAACGATGCGCTTTTGGAAATATCAT 7002
QY 1840 CTGTGCACTTACAGAGAAAGAGTTTAAATTTGGTGCATGATGAATTCGATATGT 1899
DB 7003 CCAACCATCTACTAGGGGTAAGTCTTTTAAATTTGGGAGGATGATGAATTCGGAATGT 7062
QY 1900 TCTTAACGCTGTTTGTCAACACATGATCAATATCATGATTTGCTAGCAGAGTACTAGCTG 1959
DB 7063 TCTTCACTATTTTGTCAACACAGTTTGAATGTGTTATTCGCGAGAGGTACTGAAAG 7122
QY 1960 AAGGTTAAACAGTACAGTGGCGGCGCTCTATTCGCGAGAGTAAACATATGATGATGTTG 2019
DB 7123 AGCGGCTTAAAGTTCAGATGTGCAAGCTTCAATTTGCGAGCAACACATATCATGAGAG 7182
QY 2020 TCGTCTCCGACACCTTGATGGCGAGAGATGCGCACTTGGCTGAACATGGAAGTAAAAA 2079

Db 7183 TAGTATCTGACAAAGAAATGGCTGAGAGGTGGCCACCTGGCTCAACATGAGAGTTAAGA 7242
Qy 2080 TTATGATGACGATTATTTGATCAAGACACCTACTCTGTTGGGGGATTTATCTGTGG 2139
Db 7243 TCATGAGCGCATCTGCTGAGAGACCACTTACTTCTGGCGGCGATTTATCTTGGCAG 7302
Qy 2140 ACCAGATACAGGACAGCTTGCAGAGTGGCAGACCTCTTAAAAAGGCTTTTAAAGCTTG 2199
Db 7303 ATTCGGTATCTTCCACAGCTGGCCCGTGGCGGATCCCTGAAAAGGCTGTTAAGTTGG 7362
Qy 2200 GAAACCTTTGGCAGTATGATACCCAAAGCTGGAGCCCGCGGCACTGCATGATG 2259
Db 7363 GTAAACCCCTCCACAGCCGACGAGAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 7422
Qy 2260 AAGCAATCGATGAGAAAGAAATTTGAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2319
Db 7423 AAACAAAGGCGGTGTTAGAGAGATTAACAGGACCTTTAGCACTGGCCGTGACAGACC 7482
Qy 2320 GATACGAGATCATCTGCGAGGCTGATCATCACTCTCTGTCACAGTTAGCCGAAAGCG 2379
Db 7483 GGTATGAGTACACATTTTACACCTGTCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 7542
Qy 2380 TTAAGAACTTCAAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2439
Db 7543 AAAGAGCATTCAGAGCATCAG 7602
Qy 2440 TGACGTAGTACAGCAGCACTAC-----CCACGCGGAGAGAGAGAGAGAGAGAGAGAGAG 2486
Db 7603 CAGCATATGATCTTCACTGACTGATCTGATCAACACACACACACACATGATGAGAGATCT 7662
Qy 2487 CTCAGCTGACCTTCCACAGTTTACCTTACCAATCGAGTGGCTTACAGAGATCCAAAC 2546
Db 7663 TTAACATCTGCGCGCGCGCCCTTCCCGGCGCCGACATGAGAGAGAGAGAGAGAGAGAGAG 7722
Qy 2547 CTCCTAGGCGCGCTGAGAGCGCTTGGCGCCCGCTGCTGCTCAATTCGAATCTTA 2606
Db 7723 GAGAGCAGCGCGCGCGCGCGCTGCGCCGCAAGGCGCTGCTCAATTCGAATCTTA 7782
Qy 2607 GAGAGTCAATGATCACTTCACTTCAACACAGTACCTAATCGCGCGCGAGTTCGA 2665
Db 7783 CCAAGCGCTGATGCTTCACTTCACTTCAACAGTACCTAATCGCGCGCGAGTTCGA 7842
Qy 2666 -----CCGCGCAAGAGAGAGAGAGTGTCTTAAAGCCTTAACTGAGCTTAAAAAGA 2720
Db 7843 GCCCGCCACCGCGCGCGAG 7902
Qy 2721 AGAAGCAGACAGCCAG 2777
Db 7903 AAAGCGCAG 7962
Qy 2778 TGTGTATGAGTGGAGTGGAGCA 2801
Db 7963 TGGCATTAAGTTGAGGCGCAGCA 7986

RESULT 11
US-09-507-362-101
Sequence 101, Application US/09507362
Publication No. US20030096397A1
GENERAL INFORMATION:
APPLICANT: Dubensky Jr., Thomas W.
Polo, John M.
Bell, Barbara A.
Schlesinger, Sondra
Dryga, Sergey A.
Frolov, Ilya
TITLE OF INVENTION: RECOMBINANT ALPHAVIRUS-BASED VECTORS
WITH REDUCED INHIBITION OF CELLULAR MACRO-MOLECULAR
SYNTHESIS
NUMBER OF SEQUENCES: 125
CORRESPONDENCE ADDRESS:
ADDRESS: Seed Intellectual Property Law Group PLLC
STREET: 701 Fifth Avenue, Suite 6300

CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/507,362
FILING DATE: 18-Feb-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 930049,457D6 /1196,011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 101:
SEQUENCE CHARACTERISTICS:
LENGTH: 8000 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 101:
US-09-507-362-101
Query Match 12.7%; Score 824.4; DB 11; Length 8000;
Best Local Similarity 61.9%; Pred. No. 5.4e-235;
Matches 1388; Conservative 0; Mismatches 831; Indels 25; Gaps 4;
Qy 580 ACTGACGCTATGAAGCGGAGCGCTATATTTCTCATGTGAAACAGCGCAAGTCACTTC 639
Db 5746 ACTGACTAACCGGGGTAGTGGGTACATATTTTGACGAGACAGCGCCCTGGGCACTTGC 5805
Qy 640 AACGAAATGAGTACGTCATATGTAACATACAAAGAACTATTTGATGCGGCGGCTCATG 699
Db 5806 AAAAGAAAGTCCGCTTCTGACAGACCACTTACAGAAACGACCTTGGAGACATATGCTCGG 5865
Qy 700 AGAAGTATTAAGCGCCCGCGCTGATCTGAAAGAGAGAAATGTTACAGAAAGAACTGC 759
Db 5866 AAAAGAAATTAAGCGCCCGGTGTCAGACGTCGAGAAAGAAACAACTCAACTGAGTACC 5925
Qy 760 AATTATGCGCTCTGAAGGAAATGAAAGCAGTATCAATCAAGAAAGTAAATATGA 819
Db 5926 AGATGATGCCCAAGAGCCAAAGAAAGTACGATCTGTAAGTAAAGTAAATATCAGA 5985
Qy 820 AAGCAATTAAGCGGAGAGAGCACTATTTGATGGGAGCAATCTATCATCAAGAGTGA 879
Db 5986 AAGCATTAACCACTGAGAGACTTCTGAGAGACTGACGTATTAATCTTGGCAGAGATC 6045
Qy 880 ATCTGTGAGTGTTAAGAGTCAATTAATCTGTAACCAATCTACTGCTCAACGCTAATTA 939
Db 6046 AGCCA---GAATGCTATTAAGATCACTATCCGAAACCATGTTATCTCAGTAGCTTACCGG 6102
Qy 940 ACAGCTTACATCTGACAGAGTGGCGGTAAACGTGCACTTAATTAATCAAGAGATT 999
Db 6103 CGAACTACTCGATCCACAGATTCGCTGAGCTGTCTGTAACAACTATCTGATAGAGAACT 6162
Qy 1000 ACCCTAGATGAGCAGTATTTGTTAACAAGTGAATGAGTGGATCTTGAAGAGTGGG 1059
Db 6163 ATCCGACAGTACATCTTATCAATTAAGTACGAGTACAGATGCTTACTGATATGAGAG 6222
Qy 1060 ACGCGCATGCTGTCTGATGATACAGCACTTTTGTCCGGCTAAACTGAGAGACTACC 1119
Db 6223 ACGGACAGTGGCGCTGCTGATGATGCAACCTTCTGCGCGCTAAGACTTAAGAGTTACC 6282
Qy 1120 CAAGAAAGCATGCTATTTGAGCAGCAGATAGATAGCCGCTCCATCGCTTATACAGA 1179
Db 6283 CGAAAAAATGAGTATAGAGCCCGCAATATCCGAGTGGGTTCCATCAGCGATGAGA 6342

QY	1180	ATACATTACAAAATGTAATTTGGCTGACGTACTTAAAGAAATTTGCAAGTTAACCCAAATGC	12319
Db	6343	ACAGCGCTACAAAATGTGTCTACTTCCGCACTTAAAGAAATTTGCAAGTCAACGCATATGC	6402
QY	1240	GAGAAATTACCTGTCTTAGATTGGCGCGCATTTAAATGTTGATTGTTTCAMAGAAATACCAT	1299
Db	6403	GTGAATCGCCAAACTGGACTCAGGACATTTCAATGTTCGAATGCTTTCGAAATATTCAT	6462
QY	1300	GCAATGATGACTACTGGGATACCTTTGGCGATAACCTTATTCGGCTAATCTACAGAGAAGC	1359
Db	6463	GTAAATGACGAATATTGGAGAGAGTTGCTCGGAAGCCAAATTAGAAATTAACCTAGATTTG	6522
QY	1360	TTAGGCAATATGPGACAAAGCTGAAAGGGCCGAAGACGACGACATTTGTTGCCAATCTC	1419
Db	6523	TCACCGCATATGTAAGTACGTAAGAAAGGCCCTTAAGCCGCGCAGCTATTTTCAAAAGCT	6582
QY	1420	ATAATCTTAAACCGTTGCAGAGAGATACCAATGATTCATTTGTCATGTGATCTTAAAGAG	1479
Db	6583	ATATTTTGTGTCCTCATTTGCAAGAAGTGCCTATAGATAGTTGTCATGACATGAAAAG	6642
QY	1480	ATGTCAAAGTTACTCCGGGACGAAACATACAGAGAGCGGCTTAAGGTGCAAGTTATTC	1539
Db	6643	ACGTAAAGTTAACACAGGACGAAACACAGAAAGAACCCAAAGTACAAATGATATAC	6702
QY	1540	AGGCTGACGATCCCTTGCTACCCGCTTACCTTCCGGATTCATATCCGGAATTAATCCGTA	1599
Db	6703	AAGCCGACGAACCCCTGGCGACTGTATCTTATGGGGATTCACCGGAATTAATGTGCTA	6762
QY	1600	GACTGAATGCGGTCCTTCTGCGCAATATCCATCTCTTTCGACATGTCAGCGGAAGTT	1659
Db	6763	GGCTTACGGCGGTCTTGCTTCCAAACATTCACAGCTTTTGAACATGTCCGCGGAGATTT	6822
QY	1660	TTGATGCGATTAATGCTGGAACATTTCCACACGCGGACCCAGTATTGAAACGACATCG	1719
Db	6823	TTGATGCAATCATAGACGAACATTTCAAGCAAGCGACCCGCTACTGAGACGATATTCG	6882
QY	1720	CGTGTGTTGATTAAGGCAAGACGCTATTCGCCATTTCCGCGTGTATGATCCTTGAGG	1779
Db	6883	CATATATTCGAACAAAGCCCAACACGCTATAGGCGTTAACCGGTGTGATGATCTTTGAGG	6942
QY	1780	ACTTAGGTGTGACCAACCGCTCTTAGATTGATAGAGCGCGGCTTCGGCAATATTCAT	1839
Db	6943	ACCTGGGTGTGATCAACCACTACTCGACTGTGATGATGGCCCTTTGGAAATATTCAT	7002
QY	1840	CTGTGCACCTTACCTTACAGGAACGAGGTTTAAATTTGTGCCATGATGAATTCGGTATGT	1899
Db	7003	CCACCCATCTACCTACGGGTACTCGTTTAAATTCGGGGCGATGATGAATTCGGAATGT	7062
QY	1900	TCTTAAGCGCTGTTGTCAACACACTAGTCATATCATATGATTGTCTAGCAGATATCTAGTG	1959
Db	7063	TCCTCACACTTTTGTCAACACAGTTTGAATGTCTGTTATGCGCCAGCAGATATCTAGAG	7122
QY	1960	AACGGTTAACACGTCAGCGTGGCGGCGCTCTATCGGCGACGATTAACATAGCATGGTG	2019
Db	7123	AGCGGCTTAAACGTCACGATGTGCAAGCTTCACTTGGGACGACACATCTACATGGAG	7182
QY	2020	TCGTCTCCGACACCTTGAATGCGAGAGATCGCCACTTTGCTGTAACATGGAATTAATA	2079
Db	7183	TAGATATCGACAAAGAAATGTGTAGAGGTGGCCACTGGCTCAACATGAGGGTTAAGA	7242
QY	2080	TTATTTGATGACGTTATTGGTATCAAAGACCCCTACTTGTGGGGGATTTATCCTGTGTG	2139
Db	7243	TCATTCGACGAGTCACTCGGTAGAGAACCACTTACTTGTGGGGCGAATTTATCTTGCAAG	7302
QY	2140	ACCGATTAACAGGCAACGCTGCGACAGTTCGCAAGCCCTTAAAGGCTTTTAAAGCTG	2199
Db	7303	ATTGCGTTACTTCCACACGCTGCGCGGTGGCGGATCCCTGTAAGAGCTGTTTAAAGTGG	7362
QY	2200	GAAAAACAATTCGACAGTCGATGATACCAAGAAGCTGGACCGCGCGGCGCACTGACATGATG	2259
Db	7363	GTAAACCGCTCCACGCGCAGCAGACGACGAAAGACGAAAGACGCGCTCTGCTAGATG	7422
QY	2260	AAGCAATGCGATGAAACAGAAATTTGGAATTAACGACAGCTTATGTAAGGCGGTAGATATCA	2319

Db	7423	AAACAAGGGCGTGTTAGAGTAGTATTAACAGGCACTTAGCAGTGGCCGTGACGACC	7482
Qy	2320	GATACGAGATCATACTGGCAGGCGCTGATCATCACTCTCTGTCCAGCTTAGCCGAAAGCG	2379
Db	7483	GGTATGGGTGAGCAATATTTTACACCTGTCTCTACTGGGATTGAGAACTTTGGCCAGGCA	7542
Qy	2380	TTTAAGAACTTCAAGAGCATTAAGAGGAGGAGCCCAATCACTCTCAAGCGCTGACTTAATAGG	2439
Db	7543	AAAGAGCATTTCCAAAGCATCTAGAGGGGAAATTAAGCATCTCTAGGTGTCTTAATATGT	7602
Qy	2440	TGACGTAGTACACGCACTAC-----CCACCGGAGATGTTTCCATACC	2486
Db	7603	CAGCATAGTTCATTTCACTGACTTAATACTACAACACACACCATTAATATGAGATTTCT	7662
Qy	2487	CTCAGCTGAATTTCCACCAAGTTTAACTCTCAATCAATCCGATGGCTTACCGAGATCCAAACC	2546
Db	7663	TTTAACAATCTGGCGCGCGCCCTTCCCGGCCCCCATCTGCCATGTGAGAGGCGCGGAGAA	7722
Qy	2547	CTCCTTAAGCCCGCTGAGGCGCCTTTTGGCCGCCCGCTGCTGCTCAATCGAAGATTTA	2606
Db	7723	GGAGGCAAGGGGGCCCGATCGCTGGCCGCAACGGGGCTGGCTTCTCAATCCAGCAACTGA	7782
Qy	2607	GGAGGTGATATGTCATTTGACTTTCAAACAAGATCACTTAATCCGCCCGCCAGGTCCA-	2665
Db	7783	CCACAGCGCTGATGTCCTTATGTCAATTGGAACAGGCACTAGACCTCAACCCCCCATGTCCAC	7842
Qy	2666	-----CCGCCAAGAAGAAAGAAAGATGCTCTTAAGCCAAAACTACTCAGCTTAATAAGA	2720
Db	7843	GCCCCGCAACCGCCGCAAGAGAGCAGGCGCCCAAGCAACACCGAABCCGAAGAAACCA	7902
Qy	2721	AGAAAGCAGCAAGCCAAAGAGACGAAAGC---AAGCTTAAACGAGGAAAAGCAACGTA	2777
Db	7903	AAAGCGAGGAGAAAGAAAGAACCACTGCAAAAACCCAAAACCGGAAAGAGACAGCGCA	7962
Qy	2778	TGTGTATGAAGTGTGAGATCGGACCA	2801
Db	7963	TGGCATTTAAGTTTGGAGGCGGACA	7986

```

RESULT 12
US-09-733-042-1
: Sequence 1, Application US/09733042
: Patent No. US20020168709A1
: GENERAL INFORMATION:
: APPLICANT: Hennecke, Frank
: APPLICANT: Renner, Wolfgang A.
: TITLE OF INVENTION: Replicon Based Activation of Endogenous Genes
: FILE REFERENCE: 1700.0100001
: CURRENT APPLICATION NUMBER: US/09/733.042
: PRIORITY FILING DATE: 2000-12-11
: PRIOR APPLICATION NUMBER: US 60/169,988
: NUMBER OF SEQ ID NOS: 49
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 1
: LENGTH: 11282
: TYPE: DNA
: ORGANISM: PCYtTs
US-09-733-042-1

```

	Query Match	12.6%	Score 816	DB 10	Length 11282
	Best Local Similarity	65.1%	Pred. No. 2.3e-232		
	Matches 1219	Conservative 0	Mid. Matches 650	Indels 3	Gaps 1
QY	580	ACTGACGGTATGAAAGCGGGAGCGTATATTTTCTCATCGGAAACAGGCGCAAGGTCACTTC	639		
DB	6638	ACTACTPAACCGGGGTAGTGGGTGACATATTTTGACGGACACAGGCGCTTGGCATTTGC	6697		
QY	640	AACGAATCATGTCGTCATATGTAACTACAGAACTATATTTGGATCGGGCGCTCATG	699		
DB	6698	AAAGAAGATCCGTTCTGTGACAAACGACTTACAGAACCGAAGCTTGGAGCGCATGTCTCGG	6757		

QY 700 AGAAGTATTAGCCCGCGCTCGATCTTCGAAAGAGAGAAATGTTTACAGAAAGAACTGC 759
 Db 6758 AAAAGATTATGCCCCGGGTGCTCGACAGCTCGAAAGAGAGAACTCAAACTCAGATACC 6817
 QY 760 AATTATGGCCCTCTAAGAGAAATAGAGAGGTATCATTCAGAAAGTAAAGTAAATATGA 819
 Db 6818 AGATATGCCCAACCAAGAGAGAAAGAGGTACAGCTCTGTAAGTAAAGTAAATATGA 6877
 QY 820 AAGCAATTACAGCGAGGAGCTCATTTTCGATTTGGGACATATCATCATCAGAGTGA 879
 Db 6878 AAGCATACCACTAGAGGAGCTACTGTGAGAGCTACGACTGTATTAATCTTCGACAG--- 6934
 QY 880 ATCTGTGAGTGTATACAGAGTCAATATCTGTACCAATCTACTCTGTCACGTAATTA 939
 Db 6935 ATCAGCCAGAAATGCTATTAAGATCACCTATCCGAAACATTTGACTCCAGTACGACGG 6994
 QY 940 ACAGGTTTACATCTGACAGAGCTGGGGTTAAACGTGCACTTAATTCAGAGAAAT 999
 Db 6995 CGAATCTCTCGATCCAGAGTTCGCTGTAGCTGTCTGTAAACAATCTGACATGAGAACT 7054
 QY 1000 ACCCTACAGTACGAGTATTTGTATACAGATGAAATACGATGCTATCTTGACATGTGG 1059
 Db 7055 ATCCGACAGTACATCTTATGATTAATCTGACAGATGAGATCTTATCTTGATGTGATG 7114
 QY 1060 ACAGGCGATCGTGTCTAGATACAGCCATTCTTTCGCGCTAAACTGAGAACCTACC 1119
 Db 7115 ACAGACAGTGCAGATGCTGATTAATCTGCAACCTTCTGCCCCGCTTAAGCTTGAATACC 7174
 QY 1120 CAAGAACATGATTTTGGACGAGATTAAGATCAGCCGTCCCATGCGCTTATACGA 1179
 Db 7175 CGAAAAACATGATAGATAGAGCCCGAAATATCCGAGTGGCGGTTCATCAGAGATGAGA 7234
 QY 1180 ATACATCTCAAAATGTATTTGGCTGAGCTACTTAAAGAAATGCAACGTTAACCAATGC 1239
 Db 7235 ACAGCTCAAAATGTCTCATTTGCCGCACTTAAAGAAATGCAACGTCAGCGAGATGC 7294
 QY 1240 GAGAAATTAATCTGCTTGAATTCGGGCGCAATTAATGTTGATTTTCAAGAAATACGAT 1299
 Db 7295 GTGAATCTCAACATGAGCTGAGGAGATTCATTAATGTTGATTTTCAAGAAATACGAT 7354
 QY 1300 GCAATGATGATGATGAGTACCTTTTCGCGATTAACCTTATTTGGCTTAACTACAGAAAG 1359
 Db 7355 GTAAAGACAGATGATGAGGAGATTCGCTCGGAAACCAATTAAGATTTCACTGAGTTTG 7414
 QY 1360 TTACGCAATATGTGCAAAAGCTGAAAGGCGGAAAGCAGCAGCTTGTGGCAATATC 1419
 Db 7415 TCACCGCATATGTAGCTAGCTGAAAGGCGCTTAAAGGCGCGCCTATTTTGAAGAGCT 7474
 QY 1420 ATATCTTAAACCGTTGACAGAGATACCAATGATCAATTCGTCAATGATCTTAAAGAG 1479
 Db 7475 ATAAATTTGCTCCATTTGCAAGAAAGTCCATGATGATTTGCTATGACATGAAAGAG 7534
 QY 1480 ATGTCAAAATTAATCTCCGCGACGAAACATACAGAGAGCGGCTTAAAGTGCAGTTATTC 1539
 Db 7535 ACGTGAAAGTTAACACAGGACGAAACACAGAAAGAAAGCCGAAATACAAAGTATAC 7594
 QY 1540 AGGCTGAGATCCCTTGTACCGCTTACCTTTGGGGATTCATCGGGAAATTAAGCCGTA 1599
 Db 7595 AAGCGCGAAACCCCTGGGAGCTTACTTATGCGGGAATTCACCGGGAATTAAGCGTA 7654
 QY 1600 GACTGATGCGGTCTTGTGCAAAATATCAATCTCTTGCACATGTCAAGGGAAGATT 1659
 Db 7655 GGCCTTAAGCGCGTCTTGTCTTCAAAATTCACACGCTTTTGAATGTCGCGGAGATTT 7714
 QY 1660 TTGATGCAATTAATGCTGAACATTTTCAACGCGGACCCAGATTTGAAAGGAGCATCG 1719
 Db 7715 TTGATGCAATTAATGCTGAACATTTTCAAGCAAGGCGGACCCGATCTGAGAGCGATATCG 7774
 QY 1720 CGTGTGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1779
 Db 7775 CATCTTGTGACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7834
 QY 1780 ACTTAGTGTGACCAACCGCTCTTAGATTTGATGAGGCGGCGTTCGCAATATCACAT 1839

Db 7835 ACTTGGTGTGATCAACACACTACGACTGATGAGTGGCGCTTTGAGAAATATCAT 7894
 QY 1840 CTGTGACCTTACCTACAGAAAGAGGTTAAATTTGGTCCATGATGAAATCCGATATCT 1899
 Db 7895 CCACCATCTTACCTTACGAGGATCTGTTTAAATTCGGGCGGATGATGAAATCCGAAATCT 7954
 QY 1900 TCTTAAAGCTGTTTGTCAACACACTGATCAATATCATATGCTGTGACAGTACTACGTT 1959
 Db 7955 TCTTCAACATTTTGTCAACACAGTTTGAATGTGTTATGCTCCAGAGAGTACTAGAAAG 8014
 QY 1960 AACGTTTAAACAGTCAAGCTGCGCGGCTCTATTCGCGAGATTAACATATGTCATGCTG 2019
 Db 8015 AGCGCTTAAACAGTCCAGATGTGACGCGTTTATTTGGGACGACAAATCATATCATGAG 8074
 QY 2020 TCTGTCGACACCTTGAATGCGGAGAGATGCGCCACTTTGGCTGAACATGAGAAATTA 2079
 Db 8075 TAGATCTGACAAAGAAATGCTGAGAGTGGCCACTGGCTCAACATGAGGTTAAGA 8134
 QY 2080 TTATGATGACGATTAATTTGTATCAAAAGCACTTCTGTGGGGGATTTATCTGATGG 2139
 Db 8135 TCATGAGGAGTATGATGCTGATGAGAGCACTTATCTTCTGCGGCGGATTTATCTTGCAAG 8194
 QY 2140 ACCAGATTAACAGGACAGCTGACAGAGTGGAGACCTCTTAAAGGCTTTTAAAGCTTG 2199
 Db 8195 ATTCGTTACTTCCACAGCGTGGCGGAGATCCCTGAAAGGCTGTTAAGTTGG 8254
 QY 2200 GAAAAACCTTTCGAGTATGATATCCCAAGACTGAGACCGCGCGGCGGACCTGATGATG 2259
 Db 8255 GTAAACCGCTCCAGCGCAGACAGAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 8314
 QY 2260 AAGCAATGATGAGAAAGAAATGGAATTTAGAGAGTGAAGGCGCGGAGAAATCA 2319
 Db 8315 AAACAAAGCGGTGTTTGAATGATGATTAACAGGACCTTATGACAGTGGCGGAGAGAG 8374
 QY 2320 GATACGAGATCATATGCGAGCGCTGATCATGCTCTGTCCAGCTTAAGCCGAAAGCG 2379
 Db 8375 GGTATGAGTATGACAAATTTTACACCTGTCTTACTGAGGAACTTTTGGCCAGAGCA 8434
 QY 2380 TTAAGAACTTCAAGAGCTTAAGAGGAGGCCAATCACTCTTACGCTGATCTTAATGAG 2439
 Db 8435 AAAGAGATTTCCAAAGCCATCAGAGGAGAAATTAAGCATCTTACGCTGCTTAATAGT 8494
 QY 2440 TGACGTATGAGA 2451
 Db 8495 CAGCATATGACA 8506

RESULT 13
 US-09-275-883-1
 ; Sequence 1, Application US/09275883
 ; Publication No. US20030053988A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Renner, Wolfgang A.
 ; APPLICANT: Nieba, Lars
 ; APPLICANT: Boersma, Marco
 ; TITLE OF INVENTION: Inducible Alphaviral Gene Expression System
 ; FILE REFERENCE: 1700.0020001
 ; CURRENT APPLICATION NUMBER: US/09/275.883
 ; CURRENT FILING DATE: 1999-03-25
 ; EARLIER APPLICATION NUMBER: US 60/079.562
 ; EARLIER FILING DATE: 1998-03-27
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 11282
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: cDNA
 ; US-09-275-883-1
 Query Match 12.6%; Score 816; DB 11; Length 11282;

Best Local Similarity 65.1%; Pred. No. 2,3e-222;
Matches 1219; Conservative 0; Mismatches 650; Indels 3; Gaps 1;

```
QY 580 ACTACGGATGAAACCGGAGCGTATATTTCTCATCGGAAACAGGCCAGGTACCTTC 639
Db 638 ACTACGATCAACCGGAGTGTGGGTACATATTTCTCATCGGAAACAGGCCAGGTACCTTC 6697
QY 640 AACGAAATCGATGATGATCAATGTAACCAAGAACCTATTTGGATGGGCGGTCATG 699
Db 698 AAAAGAGTCCGTTCTGCAACACAGCTTACAAACGCACTTGGAGCGCATGCTTCGG 6757
QY 700 AGAAGTATACGCGCCGCGCTCATCTCGAAGAGAAAATGTTCAGAGAACTGC 759
Db 6758 AAAAGATTCATGCCCGGTCTCGACACGTGAAAGGAAACACTCAACTCAGGTACC 6817
QY 760 AATTATGCGCTCTGAAAGAAATGAAAGCAGGTATCAATACGAAAGTAAATATGA 819
Db 6818 AGATGATGCCACCGAAGCCAAACAAAGTAGTACAGTCTCTGTAAGTAAATACGA 6877
QY 820 AAGCAATTAACGGAGGAGCTCATTTCTGATTTGGGACATCTATCATCAGAGTGA 879
Db 6878 AAGCAATTAACCGAGGAGGAGCTCATTTCTGATTTGGGACATCTATCATCAGAGTGA 6934
QY 880 ATCTGTGAGTGTAAACAGATCAATTAATCTGTACCAATCTACTGTCGCAACGGTAATTA 939
Db 6935 ATCAGCCAGATGCTATTAAGATCACTATCGAAACCACTGTACTCCAGTAGCGTACCG 6994
QY 940 ACAGGTTTACATCTGCGAGAGTCCGGTTAAACGTGCACTTGTATTCAGAGAAATT 999
Db 6995 CGAATCTCTCGATCCACAGTTCGCTGTAGCTGTGTACAACTATCTGTAGTAAACT 7054
QY 1000 ACCGTACGATGACGAGTATATGTAATACAGATGAATACGATCGATCTTACATGGTG 1059
Db 7055 ATCCGACGATGACATCTTATACAGATTAAGTACGATGATGCTTACTTGTATGTGATG 7114
QY 1060 ACAGGCGATCGTGTCTGTAGATACAGCAGCTTTTGTCCGGCTAAACCTGAGAACTACC 1119
Db 7115 ACGAGACGATGCGATGCTGTAGTACGACCTTCTGCGCCGCTAAGTTAGTACC 7174
QY 1120 CAAGAAAGCATGCTATTTTTCAGCCAGATGAATACAGCTGCTCCATCGCTTATACGA 1179
Db 7175 CGAATAAAGCATGATATGAGCCCGGATATCCGAGTGCCTGTCCATCAGCGATGCGA 7234
QY 1180 ATACATTAACAAATGTAATGCTGACAGTACTAAAGGAATGGAAGTTACCAATATGC 1239
Db 7235 ACAGCTACAAATGTAATGCTGACAGTACTAAAGGAATGGAAGTTACCAATATGC 7294
QY 1240 GAGAAATTAACCTGTCTGTAGATTCGCGGCAATTTAAATGTTGTTTCAAGAAATACGAT 1299
Db 7295 GTGAATGCGCAACACTGAGCTCAGGAGATTCATATGTCGAATGCTTTCGAAATATGAT 7354
QY 1300 GCATGATGATGACTGCGGATACCTTTGGCGAATACCTTATCGGCTAATCAAGAAACG 1359
Db 7355 GTATATGAGATATTTGGAGAGTTCGCTCGAAGCCAAATTAAGATTAACCTAAGTTTG 7414
QY 1360 TTACGCAATATGTCGCAAAAGCTGAAAGGCGGAAAGCAGAGCATTTGTTTCGATATCTC 1419
Db 7415 TCACCGCATATGTCGCAAAAGCTGAAAGGCGGCTTAAGGCGGCACTATTTTCGAAAGAT 7474
QY 1420 ATAAATCAAAAACGTTGCGAGAGATACCAATGATCAATTTGTCATGATGATTAAGAGAG 1479
Db 7475 ATAAATTTGTCCTCAATGTCAGAAAGTGCCTATGATGATGATGATGATGATGATGATG 7534
QY 1480 ATGTCAAAAGTATCTCCGCGCAGAAACATACAGAGAGCGGCTTAAGGTGAGGTATTC 1539
Db 7535 ACGTGAAGATTAACACAGGCAAGAAACACAGAAAGAAAGCCGAAAGTACAAAGTATAC 7594
QY 1540 AGGTGAGATCCCTGCTCACTCGCTTACCTTTGGGGATCCATCCGGAATTAAGCGTGA 1599
Db 7595 AAGCGCGAAGCCCTGCGAGCTGTTACTTATGAGGATTCACCGGGAATTAAGCGTGA 7654
QY 1600 GACTGAATGCGGTCTTTCGCAAAATATCAATCTCTTTCGACATGTACCGGAGATTT 1659
Db 1659 GACTGAATGCGGTCTTTCGCAAAATATCAATCTCTTTCGACATGTACCGGAGATTT 1659
```

```
Db 7655 GGCTTAAGCGCGTCTTCTCCAAACATCAACAGCTTTTTCATGTCGCGGAGATT 7714
QY 1660 TTGATGGAATTTTGTCTGAACATTTCCACAGCGGCGACCAAGTATGGAACGGACATCG 1719
Db 7715 TTGATGGAATTTTGTCTGAACATTTCCACAGCGGCGACCAAGTATGGAACGGACATCG 7774
QY 1720 CGTGTGTTGTAATAAGGGAAGAGACGCTATTCGCAATTTGGGGTGTGATGATCTTGAAG 1779
Db 7775 CATCATTTGCAAAAAGCCAGACGAGCTATGCGTTAAACCGGCTGATCATCTTGAAG 7834
QY 1780 ACTTAGTGTGACCAACCGCTCTTGAATTTGATGAGCGCGGCTTGGCAATTCAT 1839
Db 7835 ACCGTGGTGTGATCAACCACTACTGATTCGATGTCGCTTGGGAAATATCAT 7894
QY 1840 CTGTGACCTTACCTACGAGAACGAGTTAAATTTGGTGCATGATGAATCCGGTATGT 1899
Db 7895 CCACCAATCACTACGAGGATCTGCTTAAATTTGGGGGAGATGAATCCGGAATGT 7954
QY 1900 TCTTAACGCTGTTTGTCAACACACTAGTCAATATCATGATTTGCTAGCAGATCACTACGTG 1959
Db 7955 TCTTCACTTTTGTCAACACACTAGTCAATATCATGATTTGCTAGCAGATCACTACGTG 8014
QY 1960 AACGTTAACACGTCAGCGTCCGCGCTCTATCGCGCAGATTAATAGTCACTG 2019
Db 8015 AGCGCTTAAACGTCAGATGTCAGCGCTTCAATTTGGCGACGACATCATATGAG 8074
QY 2020 TCGTCTCCGACCTTGAATGCGGAGATGCGCACTTGTGGGGAATGGAATGGAATGGA 2079
Db 8075 TAGATCTGCAAAAGAAATGCTGAGAGTGCAGCTGCTCACTCACTCACTCACTCACT 8134
QY 2080 TTATTTGATGAGTATTTGATATCAAGCACTTCTGTGGGGAATTTATCTGTGTG 2139
Db 8135 TCATGACGAGTATGATGATGAGAGACACCTTACTCTGCGGGGATTTATCTGTGAG 8194
QY 2140 ACCAGATTAACGACAGCTGCAAGTGCAGACCTCTTAAAGGCTTTTAACTTG 2199
Db 8195 ATTCGTTACTTCCACAGCGCTGCGCTGCGGATCCCTGAAAGGCTCTTTAAAGTTG 8254
QY 2200 GAAACCATTTGCAAGTATGATATCCCAAGACTGCGACGCGCGCGGCACTGCATATG 2259
Db 8255 GTAAACCGCTCCAGCGAGAGAGACAGCAAGCAAGCAAGAGCGGCTCTGTAGATG 8314
QY 2260 AAGCATGATGAGAAACAGATTTGATTAAGCAAGTATGTAAGCCGATGATCA 2319
Db 8315 AAACAAAGGCTGTTTAAAGTATGATTAAGCAAGTATGTAAGCCGATGATCAAGCC 8374
QY 2320 GATACGATATATCTGCGAGCGCTGATCATCACTCTCTGTCCAGCTTAAGCGGAAGCG 2379
Db 8375 GTATGAGTATGACATATTTACACTGCTCTTACTGCGATGAGAACTTTTGGCCAGAGCA 8434
QY 2380 TTAAAGACTTCAAGAGATTAAGAGGAGGCCCATACCTCTACGCGTCACTTAATAGG 2439
Db 8435 AAAGCATTCAGAGCATGAGAGGGAATTAAGATCTCTACGGTGTCTTAATAGT 8494
QY 2440 TGACGTAGTAGA 2451
Db 8495 CAGCATGATCA 8506
```

RESULT 14
US-09-190-246-4
Sequence 4, Application US/09190246
Publication No. US20030180257A1
GENERAL INFORMATION:
APPLICANT: Parrington, Mark
APPLICANT: Klein, Michel
TITLE OF INVENTION: ALPHAVIRUS VECTORS FOR PARAMYXOVIRUS VACCINES
FILE REFERENCE: Parrington et al.
CURRENT APPLICATION NUMBER: US/09/190,246
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patent Ver. 2.0

SEQ ID NO 4
LENGTH: 8100
TYPE: DNA
ORGANISM: Semliki Forest virus
US-09-190-246-4

Query Match 11.8%; Score 764; DB 12; Length 8100;
Best Local Similarity 63.2%; Pred. No. 7.2e-217;
Matches 1175; Conservative 0; Mismatches 685; Indels 0; Gaps 0;

QY 594 GCGGAGCGTATATTTCTCATCGGAAACAGCCAGGTCACTTCAACAGAAATCAGTA 653
DB 5251 GCGGATGATATATTTCTCTCGGACACTGGCAGCATTTTCAACAAAATCCCTT 5310
QY 654 GGTCAATGTAACATCAAGAAACCTATATTGATCGGGCGCTCCATGAGAAATATACCC 713
DB 5311 AGGAGCAAACTCTCAAGTGGGACACACTGATGGGTTCACGAGAGAGAAATGATCCG 5370
QY 714 CCGGCGCTCGATCTTCGAAAGAGAGAAAATGTTACAGAAAGTCAATTAATGCGCTCT 773
DB 5371 CCAAAATTTGATTAATGAGAGGAGAGAGAGTTCCTGCTGAAATATCAGATGACCCATCG 5430
QY 774 GAGGAATTAAGACGATATCAATCAGAAAAGTAAATATGAAAGCAATTAACAGC 833
DB 5431 GAGGCTAATTAAGATGATACCAAGTCTCGAAAGTGAAGAAATGAAGCCACGGTGGTG 5490
QY 834 GAGGACATCATTTCTGGAATGGGCAATATCATCATGAGAAATCTCTGAGGT 893
DB 5491 GACAGGCTCAATCGGGGGCCAGATTGTACACGGAGGGAGATGAGCGCCATACCAACA 5550
QY 894 TACAGAGTCAATTAATCTCTGATCAATCTACTGTCAGAGTAAATTAAGGTTTACATCT 953
DB 5551 TACGGGTTGGGTTACCCCGCCCGGTACTCTCCCTACCTGATGAAAGATTTCCAAGC 5610
QY 954 GCAGAGTGGGGGTTAAACCTGCACTTAATTCAGAGAAATTAATCTTACAGTACC 1013
DB 5611 CCCGATGTAAGCAATGCGACGCGTCAACCAATCTATCCAGAAATTAATCCCAAGTGGCG 5670
QY 1014 AGTTATTGTATTAACGATGAATAGATGCGTATCTTGACATGCTGAGAGCGGCATCGTC 1073
DB 5671 TCGTACAGATTAACGATGAATAGATGCGTATCTTGACATGCTGAGAGCGGCATCGTC 5730
QY 1074 TGTCTAGATTAACGCACTTTTGTCCGGCTAACTGAGAAAGTCAACCAAGAGATAGC 1133
DB 5731 TGTCTGAGACAGGCACTTCTGCCCCGGAAGCTCCGGTGTCACTCCGAACATATATCG 5790
QY 1134 TATTTCAGCCAGATTAAGATCAGCGGTCCCATGCTTATACAGATTAATCAATTA 1193
DB 5791 TACCACTGAGCGACTGTACGACAGTCCGTCACCTTTCAAGAACACTACAGAAC 5850
QY 1194 GTATTGGCTGACGCTAATTAAGAAATTTGCAACGTTACCCAAATTCGGAATTAATCTGTC 1253
DB 5851 GTGCTAGGGGCGCCACCAAGAGAAATGCAACGTCAGCAAAATCCGAACTACCCACC 5910
QY 1254 TTAGATTGGGGGCACTTAATGTTGATTTTCAAGAAATAGCATGCAATGATGATGAC 1313
DB 5911 ATGACTGGGCAAGTTCACAGTGAAGTCTTCAAGCCCTATGCTGCTCCGAGAAATAT 5970
QY 1314 TGGGATACCTTTCCGATTAACCTTATTTGGCTTAATCAAGAAAGTAAAGCAATATGTC 1373
DB 5971 TGGGAAAGAAATATGTAACCACTTATCCGATTAACCACTGAGAAACATCACTATATGTC 6030
QY 1374 ACAAAGCTGAAAGGCGCAAGAGCAAGCACTTTGTTGGGAATCACTAATCTTAAACCG 1433
DB 6031 ACCAAATTTGAAAGGCGCAAGAGCTGCTCTTGTGCTTAAGAACCAACCTGTTGTCG 6090
QY 1434 TTGAGAGATTAACCAATGATCAATTTCTCATGATGCTTAAAGAGATGTCAAAGTACT 1493
DB 6091 CTGCAAGAGGTTCCCATGACAGATTCACGGTTCGACATGAAACGAGATCTCAAGTCACT 6150
QY 1494 CCGGACGAGAAACATTAAGAGAGGCGCTTAAGGTGAGGTTATTCAGGCTGCAATGCC 1553
DB 6151 CCAGGAGAGAAACACACAGAGAGAAAGACCCAAAGTCCAGTAAATTCAGAGCGGAGCCA 6210

QY 1554 CTGCTACCGCTTAACCTTTGGGGATCCATCGGGAATTAATCCGTAAGCTGAATCGGNG 1613
DB 6211 TTGGGACCGCTTAACCTTTGGGGATCCATCGGGAATTAATCCGTAAGCTGAATCGGNG 6270
QY 1614 CTCTGCGCAATTAACCTTTGGGGATCCATCGGGAATTAATCCGTAAGCTGAATCGGNG 1673
DB 6271 TTAGGCGCTTAACCTTTGGGGATCCATCGGGAATTAATCCGTAAGCTGAATCGGNG 6330
QY 1674 GCTGAACATTTTCCACACGCGGACCTGATTTGGAAGAGCAATCGGCTGTTGATTA 1733
DB 6331 GCTTCACTTTCCACACGCGGACCTGATTTGGAAGAGCAATCGGCTGTTGATTA 6390
QY 1734 AGCGAAGCGAGCTTATGCAATTTGCGCGTGTATGATCCCTTGAAGCTTATGATGATGAC 1793
DB 6391 AGCGAAGCGAGCTTATGCAATTTGCGCGTGTATGATCCCTTGAAGCTTATGATGATGAC 6450
QY 1794 CAACCGCTTTGATTTGATTAAGAGCGGCGTTCGCAATATCAATCTGTCACCTTACT 1853
DB 6451 CAGTACCTGCTGATTTGATTAAGAGCGGCGTTCGCAATATCAATCTGTCACCTTACT 6510
QY 1854 ACAGAAACGAGTTTAAATTTGTCCTATGATGAATTCGGTATGTTCTTAAACCTGTTT 1913
DB 6511 ACTGCAACGCTTTCAAGTTGAGCTATGATGAATTCGGTATGTTCTTAAACCTGTTT 6570
QY 1914 GTCAACACACTAGTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1973
DB 6571 ATTAACACTGTTTGAACATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 6630
QY 1974 TCAGCTGCGCGGCTTATGCGGAGATTAATCAATGATGATGATGATGATGATGATGATGAT 2033
DB 6631 TCGGCTGTCGCGGCTTATGCGGAGATTAATCAATGATGATGATGATGATGATGATGATGAT 6690
QY 2034 TTGATGCGGAGATGCGGCTTATGCGGAGATTAATCAATGATGATGATGATGATGATGAT 2093
DB 6691 CTGATGCGGAGATGCGGCTTATGCGGAGATTAATCAATGATGATGATGATGATGATGATGAT 6750
QY 2094 ATTGATTAACAGCACTTATGCGGAGATTAATCAATGATGATGATGATGATGATGATGATGAT 2153
DB 6751 ATGAGCGAAGCACTTATGCGGAGATTAATCAATGATGATGATGATGATGATGATGATGAT 6810
QY 2154 ACAGCTGCAAGAGTTCGAGACCTCTTAAAGAGCTTTTAAAGTGGAAACCAATGCGCA 2213
DB 6811 ACCGCTGCGGCTTATGAGACCTTAAAGAGCTTTTAAAGTGGAAACCAATGCGCA 6870
QY 2214 GTCATGATTAACCAAGCTGCGGAGCGGCGGCTGATGATGAAGCAATGCGATGG 2273
DB 6871 GCTGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 6930
QY 2274 AACGATTTGAATTAACGAGAGATTAAGAGCGGCTTAAAGTGGAAACCAATGCGCA 2333
DB 6931 TTCCGAGCAAGCTTGGGCGGCAAGCTGAGTGGCACTTAATGATGATGATGATGATGAT 6990
QY 2334 CTGCGAGGCTGATCAATGCTCTGTCAGCTTAAAGAGCTTTAAAGCTTCAAG 2393
DB 6991 GCGTCAAGAAATATCTCATAGCATGCGCCTTTGCGAGGAGCATTAAGCGCTTTAAG 7050
QY 2394 AGCATTAAGAGGAGCCCAATCACCTTACGCTGACCTTAATTAAGTGAAGCTGATGACA 2453
DB 7051 AATTTGAGAGGAGCTGTTATTAATCACTTACGCGGCTCTTAATTAAGTGAAGCTGATGACA 7110

RESULT 15
US-09-994-412-3
; Sequence 3, Application US/09994412
; Publication No. US20030059943A1
; GENERAL INFORMATION:
; APPLICANT: Cerna, Ulrich
; APPLICANT: Lundstrom, Kenneth
; TITLE OF INVENTION: Inhibition of Expression of a Target Gene
; FILE REFERENCE: 20787
; CURRENT APPLICATION NUMBER: US/09/994,412
; CURRENT FILING DATE: 2002-03-12

NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 10610
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: pSFV2gen(pD) vector
US-09-994-412-3

Query Match 11.8%; Score 764; DB 11; Length 10610;
Best Local Similarity 63.2%; Pred. No. 8.8e-217;
Matches 1175; Conservative 0; Mismatches 685; Indels 0; Gaps 0;

QY 594 GCGGACGCTATATTTCTCATCGGAAACAGGCCAGGTCCCTTCAACGAAATTCGTA 653
DB GCGGACGCTATATTTCTCATCGGACACTGCGACGACATTTTACCAAAAATCCGTT 5588
QY 654 CGTCAATGTAACTACAAAGAACTATATTGGATCGGCCCTCCATGAGAAATTTACGCC 713
DB AGGACGACAACTTCCAGTGGCCACAACTGATCGCGTCCAGAGAGAAATGTACCCG 5648
QY 714 CCGGCGCTCGATCTCGAAGAGAGAAATGTTACAGAAAGAACTGCATTTAGCGCCTT 773
DB CCAAAATTTGATCTGAGAGGAGAAAGCTGTGCTGAAATGACAGATCACCCATCG 5708
QY 774 GAAGAAATGAAACAGATATCATACGAAAGTAAATATGAAAGCAATTAACGCG 833
DB GAGGCTATTAAGATCTGATACAGCTCTCGCAAGTGAAGAAATGAAAGCCAGGTGTG 5768
QY 834 GAGGACCTATTTCTGATTTGGGACATATCTATCATCAGAAAGTAAATCTGTGAGTGT 893
DB GACAGCTCATATCGGGGCCAGATGTACACGGAGGAGACGTAGGCCGATACCAACA 5828
QY 894 TACAGAGTCAATTTCTGTACCAATCTACTGTCAACGTAATTAACAGGTTTACATCT 953
DB TACCGGTTCCGTTACCCCGCCCGTGTACTCCCTACCGTATCGAAAGATTTCTCAAGC 5888
QY 954 GCAAGGTCCGGCTTAAACGTCGAACTTACTTATCCAAAGAAATTAACCTACGTAGCC 1013
DB CCCGATGAGCAATCGCAGCTGTGCAAGAAATCTATCCAGAAATTAACCAAGTGTGCG 5948
QY 1014 AGTATTTGTATTAAGATGAATAGATCGATCTTTCGATGTGAGCGGCGCATCTGTC 1073
DB TCGTACCAAGATTAAGATGAATAGATCGATCTTTCGATGTGAGCGGCGCATCTGTC 6008
QY 1074 TGTCTAATACAGCCACTTTTGTCCGGCTAAACTGAGAACCTACCCAAAGACATAGC 1133
DB TGTCTTGGACAGAGCACTTCTGCCCCGGGAGAGCTCCGCTACCCGAAACATCATGCG 6068
QY 1134 TATTTGACCGCAGATTAAGATCAAGCCCTCCCATCGCTTATACGAATACATTAATAAT 1193
DB TACCAACGCGCAGCTGTAGGAGTGCCTCCCTGCTACCCCTTTCAGAACACTACAGAAC 6128
QY 1194 GTATTGGCTGAGCTACTAAAGAAATGCAAGCTTACCCAAATGCGAATTAACCTGTC 1253
DB GTGCTAGGCGGTGTCACCAAGAAATGCAAGCTTACCCAAATGCGAATTAACCTGTC 6188
QY 1254 TTAAATTCGGGCGCAATTAATGTGATGTTTCAAGAAATACGATGCAATGATGATAC 1313
DB ATGACATCGGCGAGTTCATAGTGGAGTCTTCAAGCCCTATGCTGCTCGGAGAAATAT 6248
QY 1314 TGGGATACCTTTCCGAGTAAACCTTATTTGGCTAATACAGAAAGTAAAGCAATATGTG 1373
DB TGGGAAATATGCTAAACCACTATCGGATTAACACTGAGAAATACATCACTATGTG 6308
QY 1374 ACAAGCTGAAGGCGCAAGACAGACATTTGTTGGAATTAATCATATCTAAACCG 1433
DB ACCAAATTTGAAGGCGCAAGACAGCTGCTGCTTTGCTTAAAGACCAACAATGCTTCCG 6368
QY 1434 TTGAGAGATACCAATGATCAATTCGTCAATGATCTAAAGAGATGTCAAAAGTTACT 1493
DB CTGACAGAGGTTCCCATGGAAGATTCACGGTCAATGAACGAGATGTCAAAAGTCACT 6428

QY 1494 CCGGACGAAACATACAGAGAGCGGCTTAAGGTGAGGTTATTCAGGCTGCGAGATCCC 1553
DB CAGAGACGAAACACACAGAGAAAGACCCAAAGTCCAGGTAATTCAGAGCGGAGCCA 6488
QY 1554 CTGCTACCGCTTAACCTTTGCGGAGTCCATCGGAAATTAAGCTGTAAGTCCGTTG 1613
DB TTGGCGACCGCTTAACCTGTCGCGATCCACAGGAAATTAAGAGAACTAAATGCTGTG 6548
QY 1614 CTTCGCAATATCCATCTCTCTTCGACATGTCAAGCGGAAATTTGATGATATAT 1673
DB TTACGCGCTTAACGTCACATTTGTAATGTGCGCCGAAACCTTTCAGCGCATATC 6608
QY 1674 GCTGAATTTTCCACGACGCGGACCCAGTATTTGAAACGACATCGCGTGTGATATAA 1733
DB GCTCTCACTTCCACCGAGAGACCGGTTCTAGAGCGACATGATCATTCGACAAA 6668
QY 1734 AGGAAAGCAGCGTATCGCATTTTCGCGCTGATGATTCCTTGAGACTTAAAGTGTGAC 1793
DB AGCAGAGACGACTCTTGGCTTTACAGTTTATGATCTCGAAGATCTAAGGGTGGAT 6728
QY 1794 CAACCGCTTTAATTTGATGAGGCGGCTTGCGCAATATCATCTGTGCACTTACT 1853
DB CAGTACCTGCTGACCTTGATGAGGCGCCTTTGGGAAATATCCAGCTGTCACTACCA 6788
QY 1854 ACAGGACGAGGTTTAAATTTGGTGCATGATGAAATCCGATATGTTCTTAAAGCTGTT 1913
DB ACTGACGCGGCTTCAAGTTCGAGCTATGATGAAATCGGACATGTTCTGACTTGT 6848
QY 1914 GTCACACACTAGTCAATATCATGATTTGTACAGAGTACTAGTGAAAGGTTAACACAG 1973
DB ATTAACTGTTTGAACATCACATATGCAAGCAGGTTACTGAGCAGAGCTACATGAC 6908
QY 1974 TCAAGTCCGCGCTTATTCGCGAGATTAATATGATGATGATGCTGCTCCGACACC 2033
DB TCCGCTGTCGCGCTTATTCGCGAGACAACTGTTACAGAGATGATCTCCGACAG 6968
QY 2034 TTGATGCGGAGAGATCGGCACTTGGCTGAACATGGAATTAATTTATGTCAGTT 2093
DB CTGATGCGGAGAGTCCGCTGCTGAGTCAATGAGTGAAGTATGATGAGCTGTC 7028
QY 2094 ATTGATTAAGACACCTACTTCTGTGGGGATTTATCTGTGTGACCAAGATTAACAGGC 2153
DB ATGGCGAAAAACCCCATATTTTGTGGGGATTCATATGTTTTCACAGCTCACAG 7088
QY 2154 ACAGCTTCGAGAGTTCGAGACCTCTTAAAGCTTTTAAAGCTTGAAGAAACCTTGC 2213
DB ACCGCTGCGGCTTGTTCAGACCACTTAACGCTGTTCAGAGTTGGGTAAAGCCGCTAAC 7148
QY 2214 GTGATGATACCCCAAGCTGAGACCGCGCGCGGCACTGCATGATGAAGCAATGCGATGG 2273
DB GCTGAAGACAG 7208
QY 2274 AACGAATTTGAATTTAGAGAGAGATTAAGTGAAGGCGCTGAATTCAGATACGAGATCATA 2333
DB TTCCGAGACAGGCTTGGGGCGCAACTGAGGTGGCACTAACATCTAGTATGAGGTAGAG 7268
QY 2334 CTGCGAGGCTGATCATCACTCTCTCTCAAGCTTACCGAAGCGTTAAGAACTTCAAG 2393
DB GCGTCAAAAGTATCTTCATAGCCATGACCACTTTCGAGAGGACATTAAGGCGCTTAAG 7328
QY 2394 AGCATTAAGAGGAGCCCAATCACCTTACGAGGCTGACCTAAATGAGTGAAGTATGATAGA 2453
DB AATTTGAGAGGAGCTGTTATACCTTACGCGGCTCTGATTTGCTGCTTATATACACA 7388

Search completed: November 16, 2003, 14:53:10
Job time : 1807 secs

US-08-991-840A-1

Query Match	90.7%	Score 5879.2;	DB 3;	Length 11492;
Best Local Similarity	96.2%;	Pred. No. 0;		
Matches 6239;	Conservative	0;	Mismatches 220;	Indels 29;
			Gaps	20;

QY	1	CACGCCGACGACGACCAACCGTAATATGTGAACCAACTGTGCAGAAAGCTATACCAACTC	60
Db	5012	CATGCCCAACCAACGCAACCGTAATATGTGAACCAACTGTGCAGAAAGCTATACCAACTC	50711
QY	61	CGGTGATTTCTAGAGCACCAAGTGGCGAATCACTGCTATTCGCTTGGCGCGCTCTGTTCT	120
Db	5072	CGGTGATTTCTAGAGCACCAAGTGGCGAATCACTGCTATTCGCTTGGCGCGCTCTGTTCT	51311
QY	121	CTAGCTCTGCCACACGCTCTGTCAACCGCTTGGAGGACTATATGACAGCGCGCTTTGTGGTTA	180
Db	5132	CTAGCTCTGCCACACGCTCTGTCAACCGCTTGGAGGACTATATGACAGCGCGCTTTGTGGTTA	51911
QY	181	CAGCTGATGTGCATCAAGCGAACAAGTCTACGTGGAGATCCCTGATCTCCTGGCTTGG	240
Db	5192	CAGCTGATGTGCATCAAGCGAACAAGTCTACGTGGAGATCCCTGATCTCCTGGCTTGG	52511
QY	241	ACGTCACAGCTCCCTTCTGACGTCATCTGATTCCTCACTGGAGTATTCGAAGTGCATCAGCT	300
Db	5252	ACGTCACAGCTCCCTTCTGACGTCATCTGATTCCTCACTGGAGTATTCGAAGTGCATCAGCT	53111
QY	301	TTGAAGTGTAGAACCACTCTGTACAGGACCTTACCTGGGAGTGTGTCCAAAGCTCTGTGAC	360
Db	5312	TTGAAGTGTAGAACCACTCTGTACAGGACCTTACCTGGGAGTGTGTCCAAAGCTCTGTGAC	53711
QY	361	TGGCGGAATATATGCAAGACTTTCATATCTGCTCTCTTCAGTTTCTTTCGAGCTTCACAC	420
Db	5372	TGGCGGAATATATGCAAGACTTTCATATCTGCTCTCTTCAGTTTCTTTCGAGCTTCACAC	54311
QY	421	CAGTACCGGACCAACGAGAGCGCCCATCCCTCATCTAGATCGAGCGCTTCCGACCTC	480
Db	5432	CAGTACCGGACCAACGAGAGCGCCCATCCCTCATCTAGATCGAGCGCTTCCGACCTC	54911
QY	481	CAGTTCCAAACCGCACGAGCACTAATGTACCAACCAACCAAGAGTCCCTAGAGGAGCT	540
Db	5492	CAGTTCCAAACCGCACGAGCACTAATGTACCAACCAACCAAGAGTCCCTAGAGGAGCT	55511
QY	541	CAGAAAGCGAGTTGGACGAGTACATCCGTCAACACTCCAACTGACGATATGAAGCGGAG	600
Db	5552	CAGAAAGCGAGTTGGACGAGTACATCCGTCAACACTCCAACTGACGATATGAAGCGGAG	56110
QY	601	CGTATATTTTCTCATCGGAAACAGGCGCAAGCTCACCTTCAACAGAAATCAGTACGTCAAT	660
Db	5611	CGTATATTTTCTCATCGGAAACAGGCGCAAGCTCACCTTCAACAGAAATCAGTACGTCAAT	56671
QY	661	GTAACCTCAAGAAACCTATATTTGGATCGGGCGGTCATGAGAGATTAAGCCCGCGGCG	720
Db	5668	GTAACCTCAAGAAACCTATATTTGGATCGGGCGGTCATGAGAGATTAAGCCCGCGGCG	57281
QY	721	TCGATCTCGAAGAGAGAAAATGTTTACAGAAAGAACTGCAATTTATCGGCTTC-TGAAAGA	779
Db	5727	TCGATCTCGAAGAGAGAAAGATGTTTACAGAAAGAACTGCAATTTATCGGCTTC-TGAAAGA	57861
QY	780	AATAGAACAGAGTATCAATCAACGAAAGATG- GAAATATGAAAGAAATTAACGCGGAGCG	838
Db	5787	AATAGAACAGAGTATCAATCAACGAAAGATGAAATATGAAAGAAATTAACGCGGAGCG	58481
QY	839	ACTCATTTCTGGATTGGGCACTATCTATCTATCATCAAGTGAATCTGTGAGTGTTAAG	898
Db	5847	ACTCATTTCTGGATTGGGCACTATCTATCTATCATCAAGTGAATCTGTGAGTGTTAAG	59061
QY	899	AGTCAATTATCTGTACCAATCTACTCGTCAACGGTAAATTAACAGGTTTACATCTGACGA	958
Db	5907	AGTCAATTATCTGTACCAATCTACTCGTCAACGGTAAATTAACAGGTTTACATCTGACGA	59661
QY	959	GCTGCGCGTTAAACGTCGAACCTTATGTTATTCGAAGAAATTAACCTTACAGTATGCCAGTTA	1018
Db	5967	GCTGCGCGTTAAACGTCGAACCTTATGTTATTCGAAGAAATTAACCTTACAGTATGCCAGTTA	60281

QY	1019	TTGTTAATACATGAATATCCATGCGTATCTTGAACATGATGAGACGGCGCATGCTGCTCT	1078
Db	6027	TTGCATTAACAGATGAATACGATGCGTATCTTGAATGGATGATACGGCCG -TCGTCTGT-T	6084
QY	1079	AGATTACAGCCACTTTTGTGTCCGGCTAAACTGAGAAGCTACCCAAAGAACATAGCTATTT	1138
Db	6085	AGATTACAGCCCTTTGTGTCGC - --TAACTGAGAACTACCCAAAGAAAGCTATGCTACTT	6140
QY	1139	GCAGCCAGAGATTAAGATATCGCCGTCCCATGCGCTATAACAGATATCATTACAAATGTATTT	1198
Db	6141	GCAGCCAGAGATTAAGATATGATGCTGTCCCATGCTACTATACAGAAATACATTACAAATGTATTT	6200
QY	1199	GGCTGACAGCTACTAAAAGAAATTGCAACGTTACCCCAATGGGAGAAATTAAGCTGCTTGA	1258
Db	6201	-GCTGACAGCCACTAAAAGAACTGCAACGTTACCCCAATGGGAGAAATTAAGCTGCTTGA -A	6257
QY	1259	TTTCGGCGGCATTTATGTATGTGATTGTTTCAAGAAATTAAGCATGCAATGATGATGCTGGGA	1318
Db	6258	TTTCGGCGGCATTTCAACGTTGATGTTTCAAGAAATTAAGCATGCAATGATGATGATTTGGGA	6317
QY	1319	TACCTTTGCGCATTAACCTATTTGCGCTAACTACAGAGAACGTTACGCAATATGTGACAAA	1378
Db	6318	TACCTTTGCGCATTAACCTATTTGCGCTAACTACAGAGAACGTTACGCAATATGTGACAAA	6377
QY	1379	GCTGAAAGGGCCGAAAGCAGCAGCATTTGTTGCGAATATCTATATCTTAAACCGTTGCA	1438
Db	6378	GCTGAAAGGGCCGAAAGCAGCAGCATTTATTCGCAAAACTCACAACCTTAAACCGTTGCA	6437
QY	1439	GGAAGTACCATGATGATCAATTGTCATGGAATCTTAAGAGAGATGTCAAAGTTACTCCCGG	1498
Db	6438	GGAAGTACCATGATGATGCAATTTGTATGATG -TCTTAAGAGAGATGTCAAAGTTACTCCCGG	6496
QY	1499	CACGAAACATACAGAGAGCGCGCTTAAGTGTCAGATTATTCAGGCTGCAGATGCCCTTGC	1558
Db	6497	CACAAACATATACAGAGAGCGCGCTTAAGTGTCAGATTATTCAGGCTGCAGATGCCCTTGC	6556
QY	1559	TACCGCTTACTTTGCGGAGTCCATCGGAAATTATGTCGATGACTGAATGCGGTCTCT	1618
Db	6557	TACCGCTTACTTTGCGGAGTCCATTCGAGAT -GTCCGTAGACTGAATGCTGTGCTTCT	6614
QY	1619	GCCAAATATCATCTCTCTTCGACATGTCAGCGGGAAGATTTGATGCGAATATATGCTGA	1678
Db	6615	GCCGAATATCATATCTCTCTTCGACATGTCAGCGGGAAGATTTGATGCGAATATATGCTGA	6674
QY	1679	ACATTTCCACACCGCGCACCAAGTATTTGAAACCGCATCGCGTCTGTTGTAATAACGA	1738
Db	6675	ACATTTCCACACCGCGCACCAAGTATTTGAAACCGCATCGCGTCTGTTAATAAACGA	6734
QY	1739	AGACGACGCTATCGCATTTTGGCGGTATGATCTTTGAGAGCTTAAGGTTCGACCAAC	1798
Db	6735	AGACGACGCTATCGCATTTTGGCGGTATGATCTTGAAGACCTTAAGCGTCGACCAAC	6794
QY	1799	GCTCTTGAATTGATAGAGGCGGCGTTCCGGAATATCATCTGTGACACCTACCTACAGG	1858
Db	6795	GCTCTTGAATTGATAGAGGCGGCGTTCCGGAATATCATCTGTGACACCTACCTACAGG	6854
QY	1859	AACGAGGTTAAATTTGGTGCATGATGAATCCGATATGTTCTTAAACGCTGTTTGTCAA	1918
Db	6855	AACGAGGTTAAATTTGGTGCATGATGAATCCGATATGTTCTTAAACGCTGTTTGTCAA	6914
QY	1919	CACACTAGTCAATATCATGATTTGTTAGACAGATATCTACGTGAACGTTAAACACGTCAGC	1978
Db	6915	CACACTAGTCAATATCATGATTTGTTAGACAGATATCTACGTGAACGTTAAACACGTCAGA	6974
QY	1979	GTTGCGCGGCTCTATTCGCGCACATTAACATATGTCAGATGATGTGCTTCGACACCTGAT	2038
Db	6975	GTTGCGCGGCTCTATTCGCGCACATTAACATATGTCAGATGATGTGCTTCGACACCTGAT	7032
QY	2039	GACGAGAGATAGCCCACTTGGCTGGAACATGGAAGTAAATTAATTTGATGACAGTTATTTGG	2098
Db	7033	GACGAGAGATAGCCCACTTGGCTGGAACATGGAAGTAAATTAATTTGATGACAGTTATTTGG	7092
QY	2099	TATCAAGACACCTACTTCTGTGTGGGGAATTTATCTGTGTGGAACAGATTAACGACAGC	2158

Db 7093 TATCAAGACACCTACTTCTG-CGGGAGATCATCTGTGTGATCAGATAC-GACACAGC 7150
Qy 2159 CTGAGAGTCGAGACCCCTTAAAAAGGCTTTTAAAGCTTGGAACCAATGCGCTGCA 2218
Db 7151 CTGAG-GTCGAGACCCCTTAAAAAGGCTTTTAAAGCTTGGAACCAATGCGCTGCA 7209
Qy 2219 TGATACCCMAAGTCTGACCGCGCGGCACTGATGATGAAGATGCGATGGAACAG 2278
Db 7210 TGACACCAAGACTGTGACCGTCCGCGGCACTGATGATGAAGCAATGCGATGGAACAG 7269
Qy 2279 AATTGAATTAACGAGAGTATGTAAGGCGGTAGAAATCCAGATACAGATCATCTGAC 2338
Db 7270 AATTGAATTAACGAGAGTATGTAAGGCGGTAGAAATCCAGATACAGATCATCTGAC 7329
Qy 2339 AGGCGTATCATCATCTGTCTGTCCAGCTTAGCCGAAAGCGTTAAGAACTTCAAGAGAT 2398
Db 7330 AGGCGTATCATCATCTGTCTGTCCAGCTTAGCCGAAAGCGTTAAGAACTTCAAGAGAT 7389
Qy 2399 AAGAGGAGCCCAATCACTCTACGGCTGACCTAAATAGGTAGCTAGTGAACAGCAGC 2458
Db 7390 AAGAGGAAACCAATCACTCTACGGCTGACCTAAATAGGTAGCTAGTGAACAGCAGC 7449
Qy 2459 CTACCCACCGGC-AGAAATGTTTCCATACCTCAGCTGAATTTCCACAGTTTACCTAC 2517
Db 7450 CTACCCACCGGCAGAAATGTTTCCATACCTCAGCTGAATTTCCACAGTTTACCTAC 7509
Qy 2518 AAATCCGATGCTTACCGAGATCCAAACCTCTTAGCGCGCTGAGAGCGCTTTGCGCC 2577
Db 7510 AAATCCGATGCTTACCGAGATCCAAACCTCTTAGCGCGCTGAGAGCGCTTTGCGCC 7569
Qy 2578 CCCGCTGCTGCTCAATCGAAGATCTTAGAGAGTGCATAGTCACTTGAACCTTCAACA 2637
Db 7570 CCCGCTGCTGCTCAATCGAAGATCTTAGAGAGTGCATAGTCACTTGAACCTTCAACA 7629
Qy 2638 ACGATCACTTAATCCGCGCGCAGGTCCACCGCAAGAAAGAAAGAGTGTCTTAAACC 2697
Db 7630 ACGATCACTTAATCCGCGCGCAGGTCCGCGCGAAAGAAAGAGTGTCTTAAACC 7689
Qy 2698 AAAAAGCTCTAGCTTAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGCTTAA 2757
Db 7690 AAAAAGCTCTAGCTTAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGCTTAA 7749
Qy 2758 ACCAGGAAAGCAACGATGTGTATGATGAGTGGAGTCCGACAAAGATTTCCGATCAT 2817
Db 7750 ACCAGGAAAGCAACGATGTGTATGATGAGTGGAGTCCGACAAAGATTTCCGATCAT 7809
Qy 2818 GCTGAACGCGCAAGTGAATGATATGCTGCTGTGTCGAGAGAAAGCTGATGAACCACT 2877
Db 7810 GCTGAACGCGCAAGTGAATGATATGCTGCTGTGTCGAGAGAAAGCTGATGAACCACT 7869
Qy 2878 CCACGTTGAAGAAAAATGATATGATGACCAATTAAGCGCGCTGAAATTTGAAGAGCTAG 2937
Db 7870 CCACGTTGAAGAAAAATGATATGATGACCAATTAAGCGCGCTGAAATTTGAAGAGCTAG 7929
Qy 2938 CATGTAAGACTTGAAGTACGCGAGCTTCCCAAGAACTTGAATATGACAGCTGCACTA 2997
Db 7930 CATGTAAGACTTGAAGTACGCGAGCTTCCCAAGAAATGAAATGACAGCTGCACTA 7989
Qy 2998 CACGAGCAAAACACCGGCTTCTACAACTGACACGCGCGCACTGATGAGAA 3057
Db 7990 CACGAGCAAAACACCGGCTTCTACAACTGACACGCGCGCACTGATGAGAA 8049
Qy 3058 TGGGAGATTTACCGTACGAGAGAGTGGCGGGAAGCGCAACGCGGAAGACCGATCT 3117
Db 8050 TGGGAGATTTACCGTACGAGAGAGTGGCGGGAAGCGCAACGATGGAAGACCGATCT 8109
Qy 3118 GGAACAAGAGGAGAGTGTGGCTATGTTCTAGAGAGTGCATATGAGGCAACGCGTAC 3177
Db 8110 GGAACAAGAGGAGAGTGTGGCTATGTTCTAGAGAGTGCATATGAGGCAACGCGTAC 8169
Qy 3178 GCGCTTTCAGTGTCACTTGGAAACAGAAAGGAGTGCATATGAGGATACCCCGAAG 3237
|||||

Db 8170 GCGCTTTCAGTGTCACTTGGAAACAGAAAGGAGTGCATATGAGGATACCCCGAAG 8229
Qy 3238 TTCTGAACCGTGTCACTAGTACAGCGCTTAGCGTGTTCGATGTACGTTCCATG 3297
Db 8230 TTCTGAACCGTGTCACTAGTACAGCGCTGTGTGTCTTCGATGTACATTCCTTG 8289
Qy 3298 CGAACAACACCGTGTGTCTATTCATGACCGCAAGAAAGAACTGCACTGCTCGAAGA 3357
Db 8290 CGAACAACACCGTGTGTCTATTCATGCGCCCAAGAAAGAACTGCACTGCTCGAAGA 8349
Qy 3358 GAACTGTGCAATCAATCAATTAACGACGCTGTGAGAAAGCTTTGAATTCATCAG 3417
Db 8350 GAACTGTGCAATCAATCAATTAACGACGCTGTGAGAAAGCTTTGAATTCATCAG 8409
Qy 3418 CCGGCGCAAGAGCAATTAACGATGACTTACACTGACGATCCCTACCTGCGGCTTCTG 3477
Db 8410 CCGGCGCAAGAGCAATTAACGATGACTTACACTGACGATCCCTACCTGCGGCTTCTG 8469
Qy 3478 CCCGTAATTCAGACACTCAACCGCGTGTTCAGCCCAATTAATAATGAGAACTGTGGA 3537
Db 8470 CCCGTAATTCAGACACTCAACCGCGCATGTTTAGCCCAATTAATAATGAGAACTGTGGA 8529
Qy 3538 GAAATCTGATGATGATGATGATTAAGATTCAGGTCTCGGCAATTCGCTCAATCAGC 3597
Db 8530 GAAATCTGATGATGATGATGATTAAGATTCAGGTCTCGGCAATTCGCTCAATCAGC 8589
Qy 3598 AGGCACTCGGATGTCAACCAATTCGTTACATGTCTTTGCAACCAACCACTGACATCA 3657
Db 8590 AGGCACTCGCAACGTCAACCAATTCGTTACATGTCTTTGCAACCAACCACTGACATCA 8649
Qy 3658 GGAAGACATATGAGAAATAATAGCTATCAGACATCTGGAACCTGCGCTGTGCGCA 3717
Db 8650 GGAAGACATATGAGAAATAATAGCTATTAATGATCACTGACCAATTCGCTGTGCGCA 8709
Qy 3718 CAAAGGTAATCTCTGTAGCTCAATGTCTTCAAGTGAACAGTAAACGTCAGTATCAC 3777
Db 8710 CAAAGGTAATCTCTGTAGCTCAATGTCTTCAAGTGAACAGTAAACGTCAGTATCAC 8769
Qy 3778 GAGCGAGCATCTGAAATTCATGACCGTGTGAGAAAGAAAGTGTGTGCG 3837
Db 8770 GAGCGAGCATCTGAAATTCATGACCGTGTGAGAAAGAAAGTGTGTGCG 8829
Qy 3838 TAGAGAGAGTACTTGTTCACACCGCTCATGGAAGCTGTAAAGTSCACGTTTACGA 3897
Db 8830 TAGAGAGAGTACTTGTTCACACCGCTCATGGAAGCTGTAAAGTSCACGTTTACGA 8889
Qy 3898 TCACCTGAAGAGAGCTGTGCGGCTATCAATCAATGACAGGCGCAACGCGTA 3957
Db 8890 TCACCTGAAGAGAGCTGTGCGGCTATCAATCAATGACAGGCGCAACGCGTA 8949
Qy 3958 TAACTCTATTTGAAGAGAGCTGACGCGAAAGTGTACATTAAACCACTTTCGCAAGA 4017
Db 8950 TAACTCTATTTGAAGAGAGCTGACGCGAAAGTGTACATTAAACCACTTTCGCAAGA 9009
Qy 4018 CGTCACTGAAGATGATAGTGTGCGCACTACAGCAAGGTATCGAGCAGCGCAAGAA 4077
Db 9010 CGTCACTGAAGATGATAGTGTGCGCACTACAGCAAGGTATCGAGCAGCGCAAGAA 9069
Qy 4078 GATGAACGCTGCACTAAAGCAAAACAGTGTATTCCTTACAAAGCGCAACAAAGAAATG 4137
Db 9070 GATGAACGCTGCACTAAAGCAAAACAGTGTATTCCTTACAAAGCGCAACAAAGAAATG 9129
Qy 4138 GGTCTTCAATCTGCGGATCTTATTAAGCACAAGACCACTCACTGCAAGGTAAATGCA 4197
Db 9130 GGTCTTCAATCTGCGGATCTTATTAAGCACAAGACCACTCACTGCAAGGTAAATGCA 9189
Qy 4198 CATTCAATTCGCTTGAACCGACAGTCTGCGGCTTCGTTAGCTCAGCGCTACAGT 4257
Db 9190 CATTCAATTCGCTTGAACCGACAGTCTGCGGCTTCGTTAGCTCAGCGCTACAGT 9249
Qy 4258 CACGAAGTGTTCAAAGCATCACTTCACTGACCTGACATGCAATTCGCTGAC 4317
Db 9250 CACGAAGTGTTCAAAGCATCACTTCACTGACCTGACATGCAATTCGCTGAC 9309
|||||

QY 4318 AACGAGAAATTTGGGGCTGCGAGCAGACGCAACAGCAAAATGATTAACAGGGTCTACATC 4377
| | | | |
Db 9310 AACGAGAAATTTGGGGCTGCGAGCAGACGCAACAGCAAAATGATTAACAGGGTCTACATC 9369
QY 4378 CAGGAATTTTCTGTGGGGCGAGAAAGGCTGAGTACGTATGGGGTAAACATGAAACAGT 4437
| | | | |
Db 9370 CAGGAATTTTCTGTGGGGCGAGAAAGGCTGAGTACGTATGGGGTAAACATGAAACAGT 9429
QY 4438 CAGAGTCTGGGGCCAGAGAGTGGCAACGAGGAGACCCATGAGATGGCCGATGAGATCAT 4497
| | | | |
Db 9430 CAGAGTCTGGGGCCAGAGAGTGGCAACGAGGAGACCCATGAGATGGCCGATGAGATCAT 9489
QY 4498 CATCACTATATATCATGGGACCTGAGTCTACACTGTATGTTGCTGTGTGTGCTGCT 4557
| | | | |
Db 9490 CATCACTATATATCATGGGACCTGAGTCTACACTGTATGTTGCTGTGTGTGCTGCT 9549
QY 4558 TGCTATCTGGTAGCACTGATCATGACAGCTTGATGCGCAAGAGAAAGAGACTG 4617
| | | | |
Db 9550 TGCTATCTGGTAGCACTGATCATGACAGCTTGATGCGCAAGAGAAAGAGACTG 9609
QY 4618 CTTGAGCCCATAGCCGCTTGGACCGAAGCGATACCCAGACATTAAGCGTTTGTG 4677
| | | | |
Db 9610 CTTGAGCCCATAGCCGCTTGGACCGAAGCGATACCCAGACATTAAGCGTTTGTG 9669
QY 4678 CTGCAATCGGCGCAACCAACGCTGAAACATTTGAGAACTTTGAAACATCTGTGTTTAA 4737
| | | | |
Db 9670 CTGCAATCGGCGCAACCAACGCTGAAACATTTGAGAACTTTGAAACATCTGTGTTTAA 9729
QY 4738 CAACCAACCGTTTCTGTGGGACAGTTGTGATCTCTGTGGACGCTTGTATTTCTGT 4797
| | | | |
Db 9730 CAACCAACCGTTTCTGTGGGACAGTTGTGATCTCTGTGGACGCTTGTATTTCTGT 9789
QY 4798 CCGCTGCTTTTCAATGCTGATGCTCTTTTATTTGTTGACAGGCTCTGCTGGGAAAGT 4857
| | | | |
Db 9790 CCGCTGCTTTTCAATGCTGATGCTCTTTTATTTGTTGACAGGCTCTGCTGGGAAAGT 9849
QY 4858 AGAGCGCTTGCAGATGCGACCACTGTGCGCAAAATGTTCCGGGAGTCCGCTAAAGGCGTT 4917
| | | | |
Db 9850 AGAGCGCTTGCAGATGCGACCACTGTGCGCAAAATGTTCCGGGAGTCCGCTAAAGGCGTT 9909
QY 4918 GGTGGAACGCGCAGGTTAACCGGCACTTAACCTGAGATCAACGCTGCTCATCGGAAT 4977
| | | | |
Db 9910 GGTGGAACGCGCAGGTTAACCGGCACTTAACCTGAGATCAACGCTGCTCATCGGAAT 9969
QY 4978 AACACCTTCACTAACAAGAGATACGTGACCTGCAAAATTCACACAGTCATTCCTTACC 5037
| | | | |
Db 9970 AACACCTTCACTAACAAGAGATACGTGACCTGCAAAATTCACACAGTCATTCCTTACC 10029
QY 5038 ACAAGTTAAATGCTGCGGGTCCCTCGAGTGCAGAGCATCTCAAAAGCGATTAACATG 5097
| | | | |
Db 10030 ACAAGTTAAATGCTGCGGGTCCCTCGAGTGCAGAGCATCTCAAAAGCGATTAACATG 10089
QY 5098 CCGCGTTTTTGGCGGTGTGACCTTTTCATGTGGGAGGCGCACAAATGCTTCTGTGACG 5157
| | | | |
Db 10090 CCGCGTTTTTGGCGGTGTGACCTTTTCATGTGGGAGGCGCACAAATGCTTCTGTGACG 10149
QY 5158 TGAAGAACACACACTGAGTGAAGGCGTACGATTTGCTTCAACATGCACTATATGATCA 5217
| | | | |
Db 10150 TGAAGAACACACACTGAGTGAAGGCGTACGATTTGCTTCAACATGCACTATATGATCA 10209
QY 5218 CGCAGTGCACCTAAAGTTCAACAGCTGCTGAAAGTGGGCGCGGTATAGTATCGG 5277
| | | | |
Db 10210 TGCAGTGCACCTAAAGTTCAACAGCTGCTGAAAGTGGGCGCGGTATAGTATCGG 10269
QY 5278 CAACACCAACCGCGACCTGATACGTTTGTACATGCGCTCAACGAGGTTCTCTACCGGA 5337
| | | | |
Db 10270 CAATACCAACCGCGCGCTGATACGTTTGTACATGCGCTCAACGAGTTCCTCACCGGA 10329
QY 5338 CTTGAAGTCAATAGCAGGCGCGATATCACGCGCTTTTCACTTTGACCAATAGTGTGT 5397
| | | | |
Db 10330 CTTGAAGTCAATAGCAGGCGCGATATCACGCGCTTTTCACTTTGACCAATAGTGTGT 10389

QY 5398 CATCAGAAAGGGGCTGTTTACAACTACGACTTCCCTGAGTATGAGCTATGAAACGAG 5457
| | | | |
Db 10390 CATTAAGAAAGGGGCTGTTTACAACTACGACTTCCCTGAGTATGAGCTATGAAACGAG 10449
QY 5458 AGCGTTGCGGATATTCAGCATCTCTGCTGTTATGCTACAGACATATGAGCCGCACTGA 5517
| | | | |
Db 10450 AGCGTTGCGGATATTCAGCATCTCTGCTGTTATGCTACAGACATATGAGCCGCACTGA 10509
QY 5518 CATACGCTGCTGAGGCTTCTGTCAAGAACATCCAGTCCCTTACACCCAGAGATATC 5577
| | | | |
Db 10510 CATACGCTGCTGAGGCTTCTGTCAAGAACATCCAGTCCCTTACACCCAGAGATATC 10569
QY 5578 AGGGTATGAATGTGAAGAAACAACCTCAGAGCAACCCCTGCAAGAAACAGACCATTTGG 5637
| | | | |
Db 10570 AGGGTATGAATGTGAAGAAACAACCTCAGAGCAACCCCTGCAAGAAACAGACCATTTGG 10629
QY 5638 ATGTAAATTTGAAGTGAAGCTTCTGGAGCGTCTTAATGTGCTTACGGGCAATCCCTAT 5697
| | | | |
Db 10630 ATGTAAATTTGAAGTGAAGCTTCTGGAGCGCTCTGCGAGCACTAATGTGCTTATGGGCAATCCCTAT 10689
QY 5698 CTGCAATGACATCCCTGATGACGCTTTGTGAGATCATCAGAAATCAACCAATTTTGA 5757
| | | | |
Db 10690 CTGCAATGACATCCCTGATGACGCTTTGTGAGATCATCAGAAATCAACCAATTTTGA 10749
QY 5758 AGTTAGCTGACAGTAGCAGACTGCAATTTATCTGACAGCTTTGTGTTCTCTAACAT 5817
| | | | |
Db 10750 AGTTAGCTGACAGTAGCAGACTGCAATTTATCTGACAGCTTTGTGTTCTCTAACAT 10809
QY 5818 ACAATACAAAGCTGACAGGAGGAGACATTTGCAATTCATCCCATCTCCAGACAGCTGT 5877
| | | | |
Db 10810 ACAATACAAAGCTGACAGGAGGAGACATTTGCAATTCATCCCATCTCCAGACAGCTGT 10869
QY 5878 TTTGAAGGAAGGACCAACATGATGACGCTGAGGAGAGATTAACATATTTTAGCAC 5937
| | | | |
Db 10870 TTTGAAGGAAGGACCAACATGATGACGCTGAGGAGAGATTAACATATTTTAGCAC 10926
QY 5938 ATCGAGCCCAAGCAAGCAAAATTTTATGTTTGTCTATGCGGCAAGAGTCACTGCAATGC 5997
| | | | |
Db 10927 ATCGAGCCCAAGCAAGCAAAATTTTATGTTTGTCTATGCGGCAAGAGTCACTGCAATGC 10986
QY 5998 TGAATGTAACACACCGGCGGACCAATATTTGAGAACCAATTAAGTCAACCAAGAT 6057
| | | | |
Db 10987 TGAATGTAACACACCGGCGGACCAATATTTGAGAACCAATTAAGTCAACCAAGAT 11046
QY 6058 CCAAGCGGCGGTTTCCAAACATCTTGGAGCTGAGCTGCTGCACTGTTGGGGAGGATC 6117
| | | | |
Db 11047 CCAAGCGGCGGTTTCCAAACATCTTGGAGCTGAGCTGCTGCACTGTTGGGGAGGATC 11106
QY 6118 ATCCCTCATTTGTTAGGACTTATATGTTGTCTGACGCTTATGCTTAAACACAG 6177
| | | | |
Db 11107 ATCCCTCATTTGTTAGGACTTATATGTTGTCTGACGCTTATGCTTAAACACAG 11166
QY 6178 TAGATGACTAGCGGCGGACACTGACATAGCGGT- AAAATCTGATGTAATCTTCCGAGAAC 6236
| | | | |
Db 11167 TAGATGACTAGCGGCGGACACTGACATAGCGGTAAAAAATCTGATGTAATCTTCCGAGAAC 11226
QY 6237 GTGGTGATATATGCAACGCGCGCTTGACACTAAATCTGATGTAATCTTCCGAGAAC 6296
| | | | |
Db 11227 GTGGTGATATATGCAACGCGCGCTTGACACTAAATCTGATGTAATCTTCCGAGAAC 11286
QY 6297 AGTGCATTAATGCTGTGACAGTGTCAATTAATGATATCACTACATATTAACAACACT 6356
| | | | |
Db 11287 AGTGCATTAATGCTGTGACAGTGTCAATTAATGATATCACTACATATTAACAACACT 11346
QY 6357 AATACCTTTTATGAGACTACATATGAGGCTCTAATATTAACATACATATTTTACTTAA 6416
| | | | |
Db 11347 AATACCTTTTATGAGACTACATATGAGGCTCTAATATTAACATACATATTTTACTTAA 11406
QY 6417 AATACCTATACACACTTATTAATTTCTTTATATATTTTCTTTTATTTTATTTGTTT 6476
| | | | |
Db 11407 AATACCTATACACACTTATTAATTTCTTTATATATTTTCTTTTATTTTATTTGTTT 11466
QY 6477 TAAATTTT 6484

Db 11467 TTAATAATT 11474

RESULT 2

US-08-801-263A-7

Sequence 7, Application US/08801263A

Patent No. 5811407

GENERAL INFORMATION:

APPLICANT: Johnston, Robert E.

APPLICANT: Davis, Nancy L.

APPLICANT: Simpson, Dennis A.

TITLE OF INVENTION: System for the In Vivo Delivery and

NUMBER OF SEQUENCES: 12 Expression of Heterologous Genes in the Bone Marrow

CORRESPONDENCE ADDRESS:

ADDRESS: Bell Seitzer Park & Gibson, P.A.

STREET: 1211 East Morehead Street

CITY: Charlotte

STATE: No. 5811407th Carolina

COUNTRY: USA

ZIP: 28234

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/801.263A

FILING DATE: 19-FEB-1997

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Sibley, Kenneth D.

REGISTRATION NUMBER: 31,665

REFERENCE/DOCKET NUMBER: 5470-147

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-420-2200

TELEFAX: 919-881-3175

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 11663 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-801-263A-7

Query Match 37.0%; Score 2397.6; DB 1; Length 11663;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3752; Conservative 0; Mismatches 1984; Indels 74; Gaps 9;

QY 580 ACTGACGATGTAACCGGAGCGGTATTTCTCATCGGAACAGGCCAGTCACTTC 639
Db 5707 ACTGCTAACCGGGGTAGGTGGTACATATTTTCACGACACAGGCCCTGGGCACTTGC 5766
QY 640 AACAGAAATCAGTACGTCAATGAATACTAAGAACTATATGATGGGCGCTCCATG 699
Db 5767 AAAAAGTCCGCTTGTGCAACACGCTTACGAACCGACCTTGAGCGCAATGTTCTGG 5826
QY 700 AGAATATTACGCCCCCGGCTCGATCTCGAAGAGAGAAATGTTTACAGAAAGTGC 759
Db 5827 AAAAGATCTAGCGCCCGGTGCTCGACAGTGAAGAGAGACAGCTCAAACTCAGGTACC 5886
QY 760 AATTATGCGCTCTGAAGAAATAGAGAGATATCATACGAAAGTAGAAAATATGA 819
Db 5887 AGATATGCGCCCAAGCCCAAGCAAAAGAGAGATACAGTCTGAAAGTAGAAAACCGA 5946
QY 820 AAGCAATTACAGCGGAGGAGCTATTCTGATTTGGGCAATATCTATCATAGAGTGA 879
Db 5947 AAGCATATACCTAGGAGGAGCTTTTACAGGCTACGGCTGTATTAATCTTGACAG--- 6003
QY 880 ATCTGTGAGTGTATACAGAGTCAATTATCTGTATCAATCTACTCTGTCACGTTAATTA 939

Db 6004 ATCAGCCAGAAATGCTATAGATCACTTACCCGAAACCATGATTCACAGGTATACCAG 6063
QY 940 AAGCTTAACTATCTGACAGAGTCCGGTTAAACCTGCACTTATTCAGAGAAAT 999
Db 6064 CGAATCTCTGACCCCAAGTTTGTGTAGCTGTTTGAACAACATATCTGATGAGAAAT 6123
QY 1000 ACCCTACAGTACCGAGTATGATTAACAGATGAATAGATCGTATCTTGCATGTGCG 1059
Db 6124 ACCGACGCTGACATCTTATACAGATCACGACGATACGATCTTATCTGATATGTGAG 6183
QY 1060 ACGGCGATCGTCTGTCTGATATACAGCCACTTTTGTCCGGTAACTAGAGACCTACC 1119
Db 6184 ACGGACAGTGGCTGCTGATATACGACCTTATTTGCCCCGCAAGCTTAGAATGTTACC 6243
QY 1120 CAAGAACATAGCTATTTGACAGCAGATGAATGATCCGCTCCATTCCTATATACGA 1179
Db 6244 CGAAAGACAGAGATATGAGAGCCCAACATCCGAGTGGCTTCAATCAGAGATGACGA 6303
QY 1180 ATACATTAACAAATATGATTTGCTGAGCTACTTAAAGAAATGCAACGTTACCCAAATGC 1239
Db 6304 ACAGCTTCAAAACAGCTGCTCATTTGCCGACTTAAAGAAATGCAACGTCACACAAATGC 6363
QY 1240 GAGAAATACCTGCTTATGATTTGCGGCAATTAATGTTATGTTTCAAGAAATATGCAT 1299
Db 6364 GTGAATCTCCAAACACTGAGTCAAGGACATTCACGTTGAATGCTTTGAAATATGAT 6423
QY 1300 GCAATGATGATCTGAGATACCTTTCCGCAATACCTTATTCGCTATCTACAGAAACG 1359
Db 6424 GCAATGAGAGATATTTGAGAGAGATTTGCCAAGAACCAATTAAGATCACTAGTGGTTGC 6483
QY 1360 TTACGCAATATGTCACAAAGCTGAAAGGCGGAAAGCAGACAGATTTGTTGGAATATCTC 1419
Db 6484 TTACGCAATATGTCACAAAGCTGAAAGGCGGAAAGCAGACAGATTTGTTGGAATATCTC 6543
QY 1420 ATATCTTAAACCGTTGCAAGAGATACCAATGATCAATTCGTATGATCTTAAAGAGAG 1479
Db 6544 ATATTTGCTCCATTTGCAAGAGAGATTCATGATGATGATGATGATGATGATGATGATG 6603
QY 1480 ATGTCAAGTTATCTCCGCAAGAACTATACAGAGAGAGCGCTTAAGTGCAGTTATTC 1539
Db 6604 ACGTCAAGTTATCTCCGCAAGAACTATACAGAGAGAGCGCTTAAGTGCAGTTATTC 6663
QY 1540 AGGCTGAGATCCCTGCTACCGGCTTCTTGGGAGATCCATGGGAATTAATGACCTGA 1599
Db 6664 AAGCGGCAAGACCTGCGGAGACCTGCTTACTATGCGGATTCACCGGAGTTAGTGCCTA 6723
QY 1600 GACTGAATGCGGTGCTTGTGCAATATCATCTCTTGCACATGTCAGCGGAAGAT 1659
Db 6724 GACTGAATGCGGTGCTTGTGCAATATCATCTCTTGCACATGTCAGCGGAAGAT 6783
QY 1660 TTGATGCAATATTTGCTGAACATTTCCACAGCGGAGACCGATTTGGAAGCGACATCG 1719
Db 6784 TTGATGCAATATTTGCTGAACATTTCCACAGCGGAGACCGATTTGGAAGCGACATCG 6843
QY 1720 CGTGTGTTGATTAAGAGCAAG 1779
Db 6844 CGTGTGTTGATTAAG 6903
QY 1780 ACTTATGATGAGCAACCGCTTATGATTTGATGAGAGCGGCGTTTCGCAATATCATAT 1839
Db 6904 ACTTATGATGAGCAACCGCTTATGATTTGATGAGAGCGGCGTTTCGCAATATCATAT 6963
QY 1840 CTGTGACCTTACCTACAG 1899
Db 6964 CTGTGACCTTACCTACAG 7023
QY 1900 TCTTAAAGCTGTTGTTGCAACACATGATGATGATGATGATGATGATGATGATGATGATGAT 1959
Db 7024 TCTTAAAGCTGTTGTTGCAACACATGATGATGATGATGATGATGATGATGATGATGATGAT 7083
QY 1960 AACGGTTAACAGAGTCAAGCGTGGCGGCTCTATTCGGAGAGATACATATGATGATGATG 2019
Db 7084 AACGGTTAACAGAGTCAAGCGTGGCGGCTCTATTCGGAGAGATACATATGATGATGATGAT 7143

2020 TCGTCTCGACACCTTGATGCGGAGAGATGCGCACTTGCGTGAACATGGAAATAA 2079
Db 7144 TAGATATGACAAAGAAATGCTGAGAGGTGTCGCACTGCGCTCAACATGAGGTAAAG 7203
Qy 2080 TTATTTAGCAGTATTGATATCAAGACCCCTACTCTGTGGGGATTTATCTGTGGT 2139
Db 7204 TCATTTAGCGCAGTATCGGCGAGAGACCACTTACTTCTGCGGTGATTCATCTTGCAG 7263
Qy 2140 ACCGATTAACAGGACAGCCCTGACAGTCCGAGACCCCTTAAAAAGCTTTTAAAGCTG 2139
Db 7264 ATTGGTTACTCTCAACAGCGTGTGCGGAGACCCCTTGAAGAGCTGTTTAAAGTTGG 7323
Qy 2200 GAAACATTTGCGACGATGATATCCAGACTGCGACCGCGCGGCGATGATGATG 2259
Db 7324 GTAAACCGCTCCAGCCGACGATGAGCAAGACGAAGACGAAGACGCGCTCTGCTAGATG 7383
Qy 2260 AAGCAATGCGATGGAACGAATTGGAATTACGAGCAGATTAGTGAAGGCGGTAGATCCA 2319
Db 7384 AAACAAAGGCGTGGTTTAGTAGTATTAACAGACACTTAGCAGTGGCGTGGCACTC 7443
Qy 2320 GATACGATATCACTGCGAGGCTGATCATCAGTCTCTGTCCACCTTACCGAAACG 2379
Db 7444 GGTATGAGTAGACAAATCAACCTGTCTGTGAGATTGAGAACTTTTCCAGAGCA 7503
Qy 2380 TTAGAATCTTCAAGACATAAGAGGAGCCCAATCACCCCTCAAGCGCTGACCTAAATAG 2439
Db 7504 AAAGACATTTCAAGCCATCAAGAGGGAATTAAGCATCTCTTAGGTGTCTTAATATGT 7563
Qy 2440 TGACGTAGTAGA-----CAGCGACTTACCACCGGAGATGTTTCCA 2482
Db 7564 CAGCATGATCATTTCACTGACTAATATACAAACACACCATGATGATGAGGATCTT 7623
Qy 2483 TACCTTCACTGAACTTTCCACAGTTTACCTTCAAAATCCGATGCTTACCGAGATCCA 2542
Db 7624 TTAAATGCTGCGCGCGCCCTTCCAGGCCCACTGCATGTGAGGCGCGGAGAA 7683
Qy 2543 AACCTCTTAGGCGCGCTGAGGCGGTTTGGCGCCCGCTGCTCAATGGAAGAT 2602
Db 7684 GAGAGCAGGCGGCCCGCATGCTGCTGCGCCCAATGGGTGGCTTCCCAATCCAGCACTGA 7743
Qy 2603 CTTAGAGGTGATGATCACTTCAAAAC--GATCACTTAATCGCGCCGAG 2660
Db 7744 CCACAGCGGTGAGTCCCTAGTCAATTGGAACAGCACTTAACCCCAAGCCGAC 7803
Qy 2661 GTCCACGCGCAAGAAAGAAAGAGTGTCTTAAGCCAAACCTACTAGCCTTAAAGAA 2720
Db 7804 GCCCGCGCGCGCAGAAAGAGCGGCCCAAGCAACCGAAGCGAAGAAACCA 7863
Qy 2721 AGAAGCAGCAAGCAAGAGAGCAAAAC---GCAAGCTTAACAGAGGAAAGCAACGTA 2777
Db 7864 AAACACAGGAAAGAAAGAAAGCACTTGCAAAACCAAAACCGGAAAGAGACGCTA 7923
Qy 2778 TGTGTATGAAAGTGGAGTGGCAAGACATTTCCGATC--ATGCTGAACGCGCAAGTGA 2834
Db 7924 TGGCACTTAAGTGAAGGCGCAGACTGTTGCAAGCTCAAAATGAGGACGAGATGCA 7983
Qy 2835 ATGATATGCTGCTGCTGTGCGAAGAGGCTGATGAAACCACTCACTGTTAAAGAAAA 2894
Db 7984 TCGGCGACGCACTGGCCATGGAAGGAAGGTATGAAACCACTCACTGAAAGAACTA 8043
Qy 2895 TTGATTAATGAGCAATTTAGCGCGCGTGAATTTGAAGAGGCTAGCATGATGACATTTGAGT 2954
Db 8044 TTGACACCCCTGTGTATCAAAAGCTCAAAATTCACCAAGTCTGACGATTAACGATGAGT 8103
Qy 2955 ACGGCGAGTTCGCCAGAAATGAAATCAGACGCTGACATACACGAGCAAAACCAAC 3014
Db 8104 TCGACAGTTGCGCGTCAATGAGAGTGAAGGCTTCACTTACCAAGTGAACACCTG 8163
Qy 3015 CGGGCTTCTAATCTGGGACACGCGCGGATGCTAGTATGAGAAATGGAGATTTTACCGTAC 3074
Db 8164 AAGGTTCTTACAACTGGGACCAAGGAGCGGTGATATATGAGGCGAGATTTTACCATCC 8223

Qy 3075 CGAAGAGTGTGGCGGGAAGGCGACAGCCGAAAGCCGATCTTGACAAAGAGCAGAG 3134
Db 8224 CCGCGGATGAGAGGAGGAGAGAGACAGTGTGTGCTCCGATTAATGATTAATCAAGCCGGG 8283
Qy 3135 TTGTGCTTATTTGTTATGAGAGGTGCAAATGAGGCGACCGCTTACGGCCCTTTCACTGTGCTA 3194
Db 8284 TTTGTGCGATGATGCTCTGAGAGGGCTGATAGGGAAACAAAGACCGCCCTTGGTCTGCA 8343
Qy 3195 CTTGGAACGAAGAGGGGTGACCTTAGGATACCCCGGAAGTTTGAACCGTGT--- 3251
Db 8344 CTGGAATTAGCAAGGGAAGCAATCAAGACAAACCCGGAAGGAGCAAGAAAGTGTCTG 8403
Qy 3252 -----CACTAGTTACAGCCCTATGCGTGTCTTGAATGTACAGTTCCATGCAACAAC 3305
Db 8404 CTGCAACCACTGTGACCGCCATGTGTGCTTGGAAACGTAGCTTCCATGCAATCGCC 8463
Qy 3306 CACCGTGTGTATTTCACTGACGCGAGAAAGAAACCTGACAGTGTGAAAGAAACCTCG 3365
Db 8464 CGCCCACTGCTACACCCGGAACCATCAGAGCTTCGACATCTTGAAGAGAACGTGA 8523
Qy 3366 ACAATCAAAATTACGACAGCGTGTGGAACGTTGAAAT---GTCCATCACGCGCGC 3422
Db 8524 ACCAGAGGCTTACGACACCTGTCTCAAGCCATATTCGCGTGGATTCGTCCGCGAGAA 8583
Qy 3423 CAAACGAAACATTAACCGATGACTTCAACTGACAGTCCCTACCTGGGCTTCCCGCT 3482
Db 8584 GTAAAGAGAGGTCACTAGAGAGCTTTTACCTTGACAGCCGCTACTTGGGACATGCTCGT 8643
Qy 3483 ATTGCAACACTCAACCGCGTGTTTACGCCCAATAAATTTGAGAACGTGTGGAGCAAT 3542
Db 8644 ACTGTCAACATCTGAACCGTGTCTTGAAGCCGATTAAGATGAGCAGAGTCTGGGATGAG 8703
Qy 3543 CTGATGATGATGATTAAGATACAGGTCTCGGCAAAATTCGCTTCAATCAAGCAGGCA 3602
Db 8704 CGGACGAAACACCATAGCATGACAGACTTCCGCCGATTGATGATGACAAAGCGAG 8763
Qy 3603 CTGCGATGTACCAAAATTCGTTATCACTGTCTTTCGACCAAGCATGATCAAGAG 3662
Db 8764 CAGCAAGCTCAAAATTAATATACCGTATGATGTGCTGAGAGGATCATATGTCAAAGAG 8823
Qy 3663 ACAGTATGAGAAATATGCTATGACGACATCTGGAACCTGCGCTGTCTTGGCCAAAG 3722
Db 8824 GCACATGATGATGACATGATGATGACACCTGAGGACGCTGTGAAGGCTTATGATCAAG 8883
Qy 3723 GGTACTTCTGTTAGCTCAATGTCTCCAGGTGACAGTGTAAACCGTCAATCAAGAGCG 3782
Db 8884 GATCTTTCTCTCGCAAGTGTCTTCAAGGAGCAGCTGATGACATGACATGACAGTGA 8943
Qy 3783 GAGCATGTGAAATTCATGCAACCGTGGAGAAAAAGATCAGAGAGATTTGTGCTAGAG 3842
Db 8944 GCAACTCAGCAACGTCATGCAATGACCGGCAAGATTAACCAAAATTCGTGGGACGGG 9003
Qy 3843 AGGAGTACTTGTTCACCCGCTCATGGAAGCTGTGAAGTGCACGTTTACATCACT 3902
Db 9004 AAAAATATGACCTACCTCCGTTCAACGTAAGAGATTCCTTGAACAGTATGACACCGT 9063
Qy 3903 TGAAGAGAGTGTGCGGGGATCAATACATGACAGGCGGCGGCAAGCGCTATAGT 3962
Db 9064 TGAAGAAACAAACCGCGGCTACATCACTATGCAAGGCGGAGCCGACCTATATCAT 9123
Qy 3963 CTTATCTGAGAGAGCTGACGCGAAGTGTATTAACCACTTCTGGAAGAACGTC 4022
Db 9124 CTTATCTGAGAGATATCATGAGGAAATTTACGGAAGCCACCATTCGGGAAAGAACATTA 9183
Qy 4023 CTTAGCAATGATGATGTGGGATCAACGACAGGTATCTGAGACCGGAAACGAAGTGA 4082
Db 9184 CGTACGAGTCAAGTGTGGGATTAACAGACCGGAACCGTTATGACCCGTACCGAAATCA 9243
Qy 4083 ACGGCTGCACTAAAGCAAAACATGCAATGCTTACAGAGCGAACAAAGAAATGGTCT 4142
Db 9244 CGGGTCAACCGCATCAAGCAGTGTGCTCTTAAGAGAGCAACGAAGTGGGCT 9303
Qy 4143 TCAACTGCGGATCTTATTAGGACACAGACCACTGATGTCAGATTAATTCACATTC 4202

Db 9304 TCAACTCGCCGACCTGATCAGACACGCCGACCAACGCCGCCAAGGAAATTGCACTTTGC 9363
Qy 4203 CATTCGCGTTGACACCGGACAGCTGTCGCCGTTCCGTTAGCTACACGCGCTACAGACGA 4262
Db 9364 CTTTCAAGCTATCCCGAGTACCTGATGTCCTGTTGCCACCGCGCAAGCTAGTAC 9423
Qy 4263 AGTGGTCAAAAGGATCAACCTCCGACCTGACTGCAATGCGACCAACATTTGCTGACAA 4322
Db 9424 ACGGCTTAAACATACAGCTCCAAATTAGACACAGACCACTGTGACATTGCTCACACCA 9483
Qy 4323 GAAATTTGGGCTGCGACGACGACAGAGAAATGATTTACAGGCTCTACATCCAGGA 4382
Db 9484 GGAACCTAGGGGCAAAACCCGACCAACCACTGAATGATCATCGAAACACGGTTGAA 9543
Qy 4383 ATTTTCTGTGGGGGAGAAAGGCTGAGTAAGTATGGGGTAAACCATTAACAGTCAGAG 4442
Db 9544 ACTTCAACCGTGACCGAAGATGCGCTGGAAATACATATGGGGCAATCACACAGTAAAGG 9603
Qy 4443 TCTGGGCCAGAGATCGGACACGAGCGACCAATGATGAGCCGATGATCATATCC 4502
Db 9604 TCTATGCCAAGATCTGACACAGAGACCTTCAAGATGGCCACAGAAATAGTACGC 9663
Qy 4503 ACTATATCATCGCATCCAGCTTACACTGTCAITGTGCTGTGTGTGTGCTCTTGTCTA 4562
Db 9664 ATTACTATCATCGCATCTCTGTATACACATCTTAGCCGTGACATCAGCTGTGCGGA 9723
Qy 4563 TCCGTGAGGACCTGCATCATCAGACGTTGATCCGCAAGCAAGAGAAATCGCTGA 4622
Db 9724 TGAATATGGGTAACTGTGTGACAGATATGTGCTGTAAAGCGCGCTGATAGTCCGA 9783
Qy 4623 CGCCATACGCGCTTGACACGACGCAAGCGTACCACAGACTTACGCGTTTGTGCTCA 4682
Db 9784 CGCCATATGCCCTTGCGCCCAATGCGGTGATTTCCAACTTCGCTGCACTTTTGTGCTG 9843
Qy 4683 TTGCGCCAAACCAAGCTGAAACATTTGGAGAACTTTGAACATCTGTGTTTAAACAAC 4742
Db 9844 TTAGGTGGCTAAATGCTGAAACATTTACCGAGACCATGATTAATTATGTCGAACAGCC 9903
Qy 4743 AACCGTTTCTGTGGGACAGTTGTGCACTCTTGTGGACGCGCTGTATTTCTGTTCCGCT 4802
Db 9904 AGCGGTCTTCTGTGGGTCCAGCTGTGTATACCTGTGGCGCTGTCTTCTTAATGCGCT 9963
Qy 4803 GCTTTTCACTGTCATGCTCTTTTATATGTTGTGACGCGCTGCTGCTGGGAAAGTACG 4862
Db 9964 GTTCTCATGCTGCTGCTCTTTTATAGTGTGCGGCGCTTACTCTGCGAAAGTACG 10023
Qy 4863 CTTTTCGAAATGCGACCACTGTGCCAAATGTTCCGGGATCCCGTATTAAGCGTTGTG 4922
Db 10024 CTTACGAAATGCGACCACTGTCCAAATGTGCCACAGATACCGTATTAAGCGACTTGTG 10083
Qy 4923 AACGCGAGTTTACGCGGCACTTAACTGTGAGATCAGCGTGTCTCATCGGAATTAAC 4982
Db 10084 AAAAGGAGGGGTAGCGCCCGCTCAATTTGAGATTTACTGTATGCTCTCGAGGTTTGC 10143
Qy 4983 CTTCAACTAACAGAGTAGAGGACCTGCAATTTCAACAGTCACTTCCACACAG 5042
Db 10144 CTTTCAACCAAGAGATGATTAATCTGCAATTTCAACAGTGTGCTCTCTCTTAAG 10203
Qy 5043 TTAATGTGCGGCTCCTCGAGTGCAGAGGATCCTCAAGGCGGATTTACATGCGCG 5102
Db 10204 TCAATGTGCGGCTCCTTGAATGTAGCCCGCGCTCAGCAGACTAATCCTGCAAG 10263
Qy 5103 TTTTTCGCGGTGTATCCCTTTCAATGTGGGAGGCGCAATGCTTCTGTGACAGTGA 5162
Db 10264 TCTTTGAGGGGTATCCCTTCAATGTGGGAGGAGCAAAATGTTTTCACAGTGA 10323
Qy 5163 ACACACAATGAGTGAAGGCTAGCGTGAATTCGCTCCAGCTGACCTTAAGTACAGGAG 5222
Db 10324 ACAGCAGATAGTGAAGGCTAGCGTGAATTTGATGATTTGGCGACCTGACAGCGC 10383
Qy 5223 TCGCACTAAAAGTTTCAACAGCTGTGAAAGTCCGCTGATATAGTACGCAACA 5282

Db 10384 AGCGAATTAAGTGCATATCTCCCGATGAAGTAGAGACTGCTATAGTATCGGAAACA 10443
Qy 5283 CCACCGGACCTCGATATCGTTTGTCAATGGCGTCAAGCCAGGTTCCACAGGACCTGA 5342
Db 10444 CTACAGTTTCTTAGATGTGATCGTAACGAGTCAACAGAAAGCTTAAAGCCTGA 10503
Qy 5343 AGGTATAGCAGGCGCATATCAGCGCTTTTTCACCTTTGACCATTAAGTGTCTATCA 5402
Db 10504 AAGTCATAGCTGACCAATTTACCATTTGTTTACACCATTTGATCACAAGTCCGTANCA 10563
Qy 5403 GAAAGGCGCTTTTACACTACACTTCCCTGATATGAGTGAAGTGAACAGAGCGT 5462
Db 10564 ATCCGCGCTGTGTACAACTATGACTTCCGGAATACGAGCATGAACAGAGCGT 10623
Qy 5463 TCGCGAATTTCAAGCATCCCGCTGATGTATGATACAGATGTAGCCGCACTACATAC 5522
Db 10624 TTGAGACATTTCAAGTACCTCTTGACTAGCAAGACCTCATGCGCACAGCATTA 10683
Qy 5523 GCGTGTGAAGCCTTCTGTCAAGAACTCAAGTCCCTTACACCCAGCAAGTATCAGGT 5582
Db 10684 GGTACTCAAGCTTCCGCAAGACGTGATGTCGTTACAGCAGCGCCGCTCTGTGAT 10743
Qy 5583 ATGAATGTGAGAGACAACTCAGAGCAGCCCTGCAAGAAACAGCAATTTGATGA 5642
Db 10744 TCGAATGTGAGAAACAACTCAGCGCCCACTGCAAGAAACCGCCCTTTTGGTGA 10803
Qy 5643 AAATTTGAAGTGAAGCCTGTGAGAGCTGATCTGTCTTACGGGCAATCCCTATCTGA 5702
Db 10804 AGATTTGATCAATCCCTGTGAGCGGTGATCTGATACAGGAAACATTTCCATTTCTA 10863
Qy 5703 TTGACATCCCTGATGACAGCTTTTGTGATCATCAGATCAGATCAATTTTGAAGTTA 5762
Db 10864 TTGACATCCCGAAGCTGCTCTTATACAGACATAGATGACACACTGCTCTCAACAGTCA 10923
Qy 5763 GCTGCAAGTAGCAGACTGATTTATTTCTGAGACTTGTGTGTTCTTAACTTACAGT 5822
Db 10924 AATGTATGTACAGTAGTGCATTAATCAGCGGACTTCCGAGGATGAGCTACCTGAGT 10983
Qy 5823 ACAAGCTGACAGGAGGACATTTGTCAGTCACTCCACTCAGACAGAGCTGTTTGA 5882
Db 10984 ATGTATCCGACCGGAGAGCAATGCTGTGACATTTGCAATTCAGACAGCAACCTCC 11043
Qy 5883 AGAAGGACCAACACATGTGACTCCGTAAGCAGATTAACATTAATTTAGCATGCA 5942
Db 11044 AAGGTGACAGTTCACTGATCTGTGAGAAAGAGGAGTGAAGTACATTTGACACCGGA 11103
Qy 5943 GCCCAAGCAAAATTTATAGTTTGTGATGCGGCAAGAACTCAGCTGCAATGCTGAT 6002
Db 11104 GCCCAGAGCGAACTTATGTTATGCTGTGTGTGAAGAGACAAATGCAATGCAAGAT 11163
Qy 6003 GTAAACCAAGGCGGACCACTAATTTGAGAACCAATTAAGTGCACCAAGATTTCCAG 6062
Db 11164 GCAAAACCAAGCTGATCATATGTCAGACCCCGCAAAATGACCAAAATTTCCAG 11223
Qy 6063 CGGAGTTTCCAAAACATTTGGAATGGCTGTGCACTGTTTGGGGAGCATCATCC 6122
Db 11224 CCGCATCTCAAAAACCTTCAATGAGTGGCTGTTTGTGCTTTTGGGCGCGCTGTGTC 11283
Qy 6123 TCAATGTTGAGGACTTAATAGTGTGTGTGACGCTATGCTTATTAACACAGTATGAT 6182
Db 11284 TATTAATTAATAGACTTATGATTTTGTCTGACAGATGATCTGATGACAGAGAT 11343
Qy 6183 GACTGACCGGAGACCTGA -CATAGCGGTAAATTCGATGTACTTCCGAGGAAGCGTGT 6241
Db 11344 GACCGTACGCCCAATGACCCGACAGCAAAACTGATGTACTTCCGAGGAAGCTGATGT 11403
Qy 6242 GCATATGCAACGCGCGCT-----TGACACT 6268
Db 11404 GCATATGCAATCAGGCTGTATATTAATCCCGCTTACCGCGGCAATATAGCAACACC 11463
Qy 6269 AAAACTGATGTATTTCCAGAGAAAGCAAGTCAATATGCTGTGACAGTGT 6318
Db 11464 AAAACTGAGTATTTCCAGAGAAAGCGAGTGCATATGCTGTGCGCAAGTGT 11513

RESULT 3

US-09-102-248-7

Sequence 7, Application US/09102248
Patent No. 6008035

GENERAL INFORMATION:

APPLICANT: Johnston, Robert E.

APPLICANT: Davis, Nancy L.

APPLICANT: Simpson, Dennis A.

TITLE OF INVENTION: System for the In Vivo Delivery and

NUMBER OF SEQUENCES: 12 Expression of Heterologous Genes in the Bone Marrow

CORRESPONDENCE ADDRESS:

ADDRESSEE: Bell Seltzer Park & Gibson, P.A.

STREET: 1211 East Morehead Street

CITY: Charlotte

STATE: No. 6008035ch Carolina

COUNTRY: USA

ZIP: 28234

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/102,248

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/801,263

FILING DATE: 19-FEB-1997

ATTORNEY/AGENT INFORMATION:

NAME: Sibley, Kenneth D.

REGISTRATION NUMBER: 31,665

REFERENCE/DOCKET NUMBER: 5470-147

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-420-2200

TELEFAX: 919-881-3175

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 11663 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-09-102-248-7

Query Match 37.0%; Score 2397.6; DB 3; Length 11663;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3752; Conservative 0; Mismatches 1984; Indels 74; Gaps 9;

QY 580 ACTGACGGTATGAACCGGAGCGTATATTTTCATCGGAAACGCGCAAGTCACTTC 639
DB 5707 ACTGTCMAACCGGGGTAGTGGTACATATTTTCGACGACACGCGCTGGGCACTTGC 5766
QY 640 AAGAGAAATCAGTACGTCATGTAACTACAGAAACCTATATTGGAGCGGCGCTCATG 639
DB 5767 AAAAGAAAGTCGTTCTGCAAAACCAAGCTTACAGAAACGACCTTGAGACGCAATGTTCTGG 5826
QY 700 AGAAGTATAGCGCCCGCGCTCATCTCGAAAGAGAGAAATGTTACAGAAAGAACTGC 759
DB 5827 AAGAAATCTACGCCCCCGGTCTGACACGTCGAAAGAGAAACAGCTCAAACTCAGGTACC 5886
QY 760 AATTATGCGCTCTGAAGAAATGAAGAGTATCAATCAAGAAAGTATGA 819
DB 5887 AGATGATGCGCCACGAGCAACAAAGAGAGTACAGTCTCGAAAGTATGAAGAACAGAG 5946
QY 820 AAGCAATTAAGCGGAGCGATCTATTCTGATTTGGCAGATATCTATCAATCAGAAATGA 879
DB 5947 AAGCCATTAACGACCTGAGCGAGTCTTTCAAGGCGTACGAGCTGTATTAACCTGCCACAG--- 6003
QY 880 ATCCTGTGAGTGTACAGAGTCAATTATCTGTACCAATCTACTCTCAACGGTATTA 939

DB 6004 ATCAGCAGAAATCTATAAGATCACTACCGCAACCATGTATTCAGAGAGTACAG 6063
QY 940 ACAGGTTTACATCTGACAGAGGTGCGGTTAAACGTGCAACTTATATCCAGAAATT 999
DB 6064 CGAAGTACTTGAACCCAAAGTTTGTGAGTCTGTGTATTAACAATCTGTCAGAGAAATT 6123
QY 1000 ACCCTCAGTACGAGTATTTATATATAGATGATATACGATGAGTATCTGACATGATG 1059
DB 6124 ACCCGACGATGATCTTTATCAATATCAGACGATGATGATCTTATGATATGATG 6183
QY 1060 ACCGCGCATGCTGCTGTATGATACAGCACTTTTGTCCGAGTAACTGAGAACTACC 1119
DB 6184 ACGGACAGTCCGCTGTGATATGCAACTTTTGTCCCGCAAGCTTAAAGATTACC 6243
QY 1120 CAAAGAGCTATGCTATTTGACCGCAGAGATTAAGATCAGCCGCTCCATGCTATACAG 1179
DB 6244 CGAAAGACACGAGTATAGAGCCCAACATCCGAGTGGGTTCCATACAGCATGACAG 6303
QY 1180 ATACATTTAACAATGTATTGGCTGACAGCTACTTAAAGAAATTGCAAGTTACCAATGC 1239
DB 6304 ACAGTTTGCAAAACGTGCTCATTGTCGCGACTTAAAGAACTCAACGTACCAAAATGC 6363
QY 1240 GAGAAATTAACCTGCTTATGATTCGCGGCACTTTATGTTGATTTCAAGAAATACGAT 1299
DB 6364 GTGACCTGCCAACACTGAGCTCAGCAGCAATTCAGCTGATGATCTTTCGAAATATGAT 6423
QY 1300 GCATGATGATGATCTGGGATACCTTTCCGATACCTTATCCGCTATCTACAGAGAACG 1359
DB 6424 GCAATGACGAGTATTTGGAGAGTTTCCGAAAGCAATTAAGATCAGTACAGAGTTG 6483
QY 1360 TTACGCAATATGATGACAAAGCTGAAAGGCGCAACAGCAGATGTTTGGCAATATCTC 1419
DB 6484 TTACCGCATRGTGCGCAAGCTTAAAGGCTTAAAGGCGCGCACTGTTGCAAAAGCGC 6543
QY 1420 ATATATTAACCGGTTGACAGAGATACCAATGATCAATTTGCTATGATCTTAAAGAG 1479
DB 6544 ATATATTTGGTCCATTGCAAGAGTCCATATGATGATGATGATGATGATGATGATGAT 6603
QY 1480 ATGTCAAGTATATCCCGGACAGAAACATACAGAGAGCGGCTTAAGTGCAGATTATTC 1539
DB 6604 ACCTGAAAGTTACACCTGGACAGAAACAGAAAGAAAGCAAGTATGATGATGATGATG 6663
QY 1540 AGGCTGAGATCCCTGCTGACCGCTTACCTTGGCGGATCCATCGGGAATGATGCGGAT 1599
DB 6664 AAGCCGCAAGCCCTGCGGACCGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 6723
QY 1600 GACTGATGCGGTGCTTGTGCAAAATATCATATCTCTTTCGACATGTCAGCGGAAGATT 1659
DB 6724 GCTTACAGCGTTTGTGTAACCAATTCACACGCTCTTGTGACATGTGCGCGGAGGACT 6783
QY 1660 TTGATGCGATATGCTGTAACATTTCCACACGCGGACCCAGTATGGAACCGGACATGC 1719
DB 6784 TTGATGCAATCAATGAGAGAACATTTCAACGAGAGTACCCGAGTACGAGAGCGGAATGC 6843
QY 1720 CGTGTGTTGTAAGAGAGAGAGAGAGTATGCGCATTTTCGAGGCTGATGATCTTGAAG 1779
DB 6844 CTTGTTGCAAAAGCCCAAG 6903
QY 1780 ACTTATGATGACCAACCGCTTATGATTTGATAGAGCGGCTTCCGCAATATCAAT 1839
DB 6904 ACTTGTGATGACCAACCACTACTGACTGATGATGATGATGATGATGATGATGATGAT 6963
QY 1840 CTGTGACCTACTTACAG 1899
DB 6964 CCAACCATCTGCGCAAG 7023
QY 1900 TCTTAAAGCTGTTTGTGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1959
DB 7024 TCTTCAAGCTCTTTTGTGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7083
QY 1960 AACGTTAAACAG 2019

Dh 7084 AGCGCTTAAACGTCCTCAATGTGACGATTTATCGGCGACGACAATTATACAGGAG 7143
Qy 2020 TCGTCTCCGACACCTTGATGGCGAGAGATGCGCCACTTGATGAAATGAAATGAAA 2079
Db 7144 TAGTATCTGACAAAGAAATGCTGAGAGGTGTGCCACTTGCTCAACATGAGGTTAAGA 7203
Qy 2080 TTATGATGCAATTTGGTATCAAGACCTTACTTCTGTGGGGATTTATCTGTGG 2139
Db 7204 TCATGACGAGATCGGCGAGAGACACTTACTTCTGTGGGTGATTCATCTTGCAAG 7263
Qy 2140 ACCAATATACAGGACAGCTCTGAGAGTGTGACAGCCCTCTTAAAGGCTTTTAAAGCTTG 2199
Db 7264 ATTCGGTAACTCCACAGGCTGTGCGGCGACCCCTTGAAGGCTGTTTAAAGTTGG 7323
Qy 2200 GAAACCAATGCGATCGATGATACCAAGACTGAGCCGCGCGGCACTGCAATGATG 2259
Db 7324 GTAAACCGCTCCAGCGAGATGAGCAAGAGAGAAAGACGCGCTGTGCTAGATG 7383
Qy 2260 AAGCAATGCGATGAGAACAGAAATTTGAAATTTACGAGAGTTAGTGAAGCCGTAGAAATCA 2319
Db 7384 AAACAAAGCGGTGTTAGAGATGATTAACAGACACTTACAGTGGCGGTGGCACTC 7443
Qy 2320 GATAGAGATCACTACGCGAGCGCTGATCATCGTCTGTGCAAGTTAGCCGAAAGCG 2379
Db 7444 GGTATGAGGTAGACAAACATCACACCTGTCTGTGCAATGAGAACTTTGGCCAGAGCA 7503
Qy 2380 TTAAGACCTTAAAGACATTAAGAGGAGCCCAATGACCTCTACCGCTGACTTAATAG 2439
Db 7504 AAAGGCACTTTTAAAGCATCAAGAGGAGAAATTAAGCATCTCTACGCTGTCTTAATAGT 7563
Qy 2440 TGACGTAGTAGA-----CACGCACTTACCCGACCGGCGAGATGTTTCCA 2482
Db 7564 CAGCATATGATCACTTCACTGACTAATACCAACAACACACCACTGAATAGAGATCT 7623
Qy 2483 TACCTCAGCTGAATTTTCCACAGTTTACCTTAACAATCGATGCTTACCGAGATCCA 2542
Db 7624 TTAACATGCTCGCGCGCGCCCTTCCAGCCCACTGTCATGAGGCGCGCGAGAA 7683
Qy 2543 AACCTCTAGGCGCGCGTGGAGGCGTTTGGCCCGCTGGCTGCTCAATCGAAGAT 2602
Db 7684 GGAGCAGGCGGCGCGGATGCTGCGCCGCAATGGCTGGCTTCCAAATCCGCAACTGA 7743
Qy 2603 CTTAGAGAGTCAATAGTCAACTTGAATTTCAAAAC--GATCACTAATCCGCGCGAG 2660
Db 7744 CCAGCGCGTCACTGCTCCCTTATGATTTGACAGGCACTAGACTTAAACCCACCCAC 7803
Qy 2661 GTCCACCGCGAAAGAAAGAGAGTGTCTTACGCCAAACCTTACGCTTAAAGAA 2720
Db 7804 GCCCGCGCGCGCGAGAAAGAGGCGCCAAACCAACCAAGCGAGAAACCA 7863
Qy 2721 AGAAGCAGCAAGCCAAAGAGAGAGAAAC---GCAAGCTTAACAGAGAAAGCAAGT 2777
Db 7864 AAACACAGAGAGAAAGAAAGAAAGCACTGCAAAACCCAAACCGGAAAGAACGCT 7923
Qy 2778 TGTGTATGAAGTTGGATCGAGACATTTCCGATC---ATGCTGAACGCGCAAGTGA 2834
Db 7924 TGGCACTTAAGTTGAGGCGCGACAGACTGTTGCAAGTCAAAAATAGAGAGAGATGTCA 7983
Qy 2835 ATGATATGCTGCTGTGTGCGAGAGGCTGATGAACCACTTCCAGCTTGAAGGAAAA 2894
Db 7984 TCGGCGACGCACTGCGCATGAGAAAGGTATGAACCACTCCACGTAAGAAAGAACT 8043
Qy 2895 TTGATTAAGAGCAATTAAGCGCGCTGAAATTTGAAGAGGCTAGCATGTCAGCTTGGAGT 2954
Db 8044 TTGACACCGCTGTGTATCAAACTCAAAATTCACCAAGTGTGACATGAGAGT 8103
Qy 2955 ACGGAGCGTTTCCAGAAATGAAATCAGACCGCTGCAATACACAGCAAGAACAC 3014
Db 8104 TCGCAACGTTGCGGTCAACATGAGAGAGGCGCTTACCTTACACAGTGAACACCTTG 8163
Qy 3015 CGGCGCTTCTACAACTGGCACCAAGCGCGAGTCCAGATGAGAGATGAGATTTACCGTAC 3074
Db 8164 AAGGCTTCTACAACTGGCACCAAGCGCGGTGAGATATGAGAGGAGATTTACATCC 8223

Qy 3075 CGAGAGAGTGGCGGAGAAAGGCGACAGCGGAAGACCGATCTTGGACAACAGAGCGAG 3134
Db 8224 CCCCGAGAGTGAAGAGAGAGAGACAGTGTGTCTGCAATATGATTAATCTACAGCGCG 8283
Qy 3135 TTGTGCTAATTTCTAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3194
Db 8284 TTGTGCGATAGTCTCGGAGAGGCGGTATGAGAGAGAAAGAAAGAGAGAGAGAGAG 8343
Qy 3195 CTGGAACCAAG 3251
Db 8344 CCGGAATAGCAAG 8403
Qy 3252 -----CACTAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3305
Db 8404 CTGACACACTGATGACAGGCGATGTGCTTGTGGAAGAGAGAGAGAGAGAGAGAGAG 8463
Qy 3306 CACCGTGTGCTAATTCATGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3365
Db 8464 CGCCACATGCTACACCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8523
Qy 3366 ACAATCCAAATTAAGACAGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3422
Db 8524 ACCAGAGCGCTAGACAG 8583
Qy 3423 CCAAGAGAGAGATTAAG 3482
Db 8584 GTAAAG 8643
Qy 3483 ATTGAG 3542
Db 8644 ACTGACACATATGAG 8703
Qy 3543 CTGATGATGATCGATTAAGATTCAGAGTCTGCGCAATTTGGCTACATAGGAGAGCA 3602
Db 8704 CGAG 8763
Qy 3603 CTGCGAG 3662
Db 8764 CAGCAAGTCAATTAAGTACCGCTTACATGCTGCTGAGAGAGAGAGAGAGAGAGAGAG 8823
Qy 3663 ACAGATGAG 3722
Db 8824 GCACATGATGATGATCAAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8883
Qy 3723 GGTACTTCTGTAGCTCAATGTCTTCCAGAGTGAAGTGAACCTGATACAGAGAGAG 3782
Db 8884 GATACTTCTCTGCGAGAGAGTCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8943
Qy 3783 GAGCATCTGAATTAATGAG 3842
Db 8944 GCAACTGAG 9003
Qy 3843 AGAGATCTTGTTCACACCGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3902
Db 9004 AAAAATATGAG 9063
Qy 3903 TGAAG 3962
Db 9064 TGAAG 9123
Qy 3963 CCTATCTGAG 4022
Db 9124 CCTATCTGAG 9183
Qy 4023 CCTATGAG 4082
Db 9184 CGTAG 9243
Qy 4083 ACGGTGACATTAAG 4142
Db 9244 CGGCTGACCGCATCAAG 9303

QY	4143	CCAACTGCGCGGATCTTATTATGAGACACAGACCACTGAGTCGAAGTAAATTGCACATTTC	4202
Db	9304	TCAACTGCGCGGATCTGATCGACACGCGCACACAGCGCCCAAGGGAATAATTGATTTGC	9363
QY	4203	CATTCCGCTTACACCGCAAGCTGTGCCGGTTCGGTTAGCTCACACGCTTACAGTCAAGA	4262
Db	9364	CTTTTCAGAGTATCCCGAGTACTCTGATGTGCTCTGTTCGCCACGCGCCGAGCTAATAC	9423
QY	4263	AGTGGTTCAAAAGGATCACTCTCACTGACTGCAATGCGACCAACATTTGTCGACACGA	4322
Db	9424	ACGGCTTTAAACACATCAAGCTCCCAATTTAGACAGACCAATCTGACATTTGCTCACACCA	9483
QY	4323	GAAATTTGGGGCTGCGAGCAGACCCACACAGAGAAATGATTTACAGGGCTTACATCCAGA	4382
Db	9484	GGAGACTAGGGGGCAAAACCCGGAACTAACACTGTAATGGATCATGTGAAACACGGTTAGAA	9543
QY	4383	ATTTTTCGTGGGGCGAGAAAGGGCTGAGTACGTATGGGGTAAACCATGAACCAAGTCAGAG	4442
Db	9544	ACTTCACCGCTGCACCGAGATGCGCTGGAAATCATATGGGGCAATCAAGAACAGTAAGGG	9603
QY	4443	TCCTGGGCCGAGAGTCCGACACAGCGCACCCCACTGAGATGGCCGATGAGATCATATCC	4502
Db	9604	TCATATGCCCAAGAGTGTGCACACAGAGACCTCTCAGGATGGCCACAGAAATAGTACAGC	9663
QY	4503	ACTATTAATCAATCCGCAATTCAGTCTACATGTGCATTTGTGCTGTGTGTGTGCTCTTCTA	4562
Db	9664	ATTAATCAATCCGCAATTCCTGTGTGTACCACTTAAGCCGTGCATCAGCTGTGTGCGCA	9723
QY	4563	TCCTAGTAGGACACTGCATCATCACAGCTTGCATTCGCCAAAGCAAGAAAGACTGCTGA	4622
Db	9724	TGATGATTTGGGCTATCTGTTCGACACTTAATGTGCTGTAAAGCGCCGCTGAGTCTGA	9783
QY	4623	CGCCATACGCGCTTGACCGAAGCGCAACGTAACCAAGACTTAAGCGGTTTGTGCTGCA	4682
Db	9784	CGCCATATGCGCTGACCGCAATATCCGATTTCCAACTTGCTGCTGCACTTTTGTGCTGTG	9843
QY	4683	TTGCGGCCAACCAAGCTTAAACATTTTGGAGAACTTTGAACCATCTGTGGTTTAAACAAC	4742
Db	9844	TTAGTGTGGCTAAATGTCTGMAACATTTCAACGAGACCATATGATTTATTTGTGTGCAAGCC	9903
QY	4743	AACGGTTTCCTGGGACAGATTGTGCAATCTCTCTGGGACGCGCTTGAATTTCTGTTCCGCT	4802
Db	9904	AGCGCTTCTTGTGGGTCCAGCTGTGTATACCTCTGGCCGCTGTGTGTCTTAATGCGCT	9963
QY	4803	GCTTTTCATGCTGCATGCTTTTATTTATTTGTTGAGGCGCTCTGCCGTGGGAAAGGTAGCG	4862
Db	9964	GTTGCTATGCTGCTGCTGCTTTTATTTAGTGTTCGCGGCGCTTACTCTGGCAAGGTACAG	1002
QY	4863	CCTTGGAACATGCGACCACTGTGCAAAATGTTCCGGGGATTCCTGATTAAGCGTTGTGCTG	4922
Db	10024	CCTAGGAACATGCGACCACTGTGTTCCAAATGTGCCACAGATACCTGATTAAGGCATCTGTG	1008
QY	4923	AACGCGCAGGTTAAGCGCACTTAAACCTGAGATCAAGGTGTCTCATTCGGAATTAACAC	4982
Db	10084	AAAGGGCAGGTTAAGCGCCCGCTCAATTTTGAAGATTACTGTATATCTCTCGAGGTTTGC	1014
QY	4983	CTTCAACTTAACAGAGATACGTGACCTTGCAAAATTCACACAGTATTCCTTTACACCAAG	5042
Db	10144	CTTCAACCAACCAAGATATTAACCTGCAAAATTCACACATCTGTGTGCTCCCTCCCTAAAG	1020
QY	5043	TTAAATGCTGGGCTCCTCGAGTGCAAAGGCAATCTTCAAAAGCGGATTAACATGCGCG	5102
Db	10204	TCAGATGCTGGGCTCCTTGTGAATGTACGCGCGCGCTCACGCGAGACTATCTGCAAGG	1026
QY	5103	TTTTTTGGGGGTGTACCTTTCAATGTGGGGAAGCGCAATGCTTCTGTGACAGTAGA	5162
Db	10264	TCCTTTGAGGGGTGTACCCCTTCAATGTGGGGAAGGACCAATGTTTTTGTGCACGTAGA	1032
QY	5163	ACACACAAGTATGAGGCGGTATGCTGAGTTGCTGCTCAGACTGCATTAATGATATCCGAG	5222
Db	10324	ACAGCGAGTATGATGAGGCGGTATGCTGAAATTTGTCAATGATTTGGCGACATGACACGCGC	1038
QY	5223	TGCGACTTAAAGTTCAACAGCTGCTTGAAGTGTGCGCTGCTGATTAATGATTAACGGCAACA	5282

Db	10384	AGGCGATTAAAGTGGCATACTGCCGCGATGAAAGTASGACTGCCGTATAGTGTACGGGAACA	10444
Qy	5283	CCACCGCGCACCTTGATACGTTTGTCAATGGCGTCACGCAGGTTCTCACGGACCTGA	5342
Db	10444	CTACCAAGTTTCTTAAGATGTGTACGTGAACGGAATGCACACAGGAACGTCTAAAGACTCTGA	10502
Qy	5343	AGGTCAATGACAGGCGCGATATCAAGCCGCTTTTCAACCTTTTGACCTTAAGTGTGTATCA	5402
Db	10504	AAGTCATAGCTGGAACCAATTTCAGCATGTGTTTACACCATTTGATCAAGAGTGTGTATCA	10563
Qy	5403	GAAAGGGCGCTTGTTTCAACTGACACTCCCTGATATGAGTATGAAGACAGAGCGGT	5462
Db	10564	ATCGCGGCTCTGGTGTACACTTAATGACTTTCCGGAATACGAGCGATGAAACAGAGCGGT	10622
Qy	5463	TCGGCGATTTTCAAGCATCTCTCGCTTGATGCTACAGACATAGTAGCCCGCATGACATAC	5522
Db	10624	TTGGAGACATTCAAGTACTCTCTTGACTAGCAAAAGCCATCTCGCAGACAGACATTTA	10683
Qy	5523	GGCTGCTGAAGCCTTCTGTCAAGAACATCAAGTCCCTTACACCCAGCATATCAAGGT	5582
Db	10684	GGCTACTCAAGCCTTCCGCAAAAGAGTGAATCCCGTATCAACGACGCGCATCTGAGAT	10744
Qy	5583	ATGGAATGTGGAAGAACATCACTGAGAGACCCCTGCAAGAAACAGACCATTTGATATGA	5642
Db	10744	TCGGAATGTGGAAGAAACACTCAAGGCGCGCCACTGACAGGAACCGCCCTTTTGGGTGCA	10802
Qy	5643	AAATTGAAATGTGAGCCTCTGCGAGCGCTTAACTGTGCTTACGGGACATCTCCATCTCGA	5702
Db	10804	AGATTGCAAGTCATTCGCGCTTTCGAGCGGTGGAAGTGTCTATACGGGAACATTTCCATTCTTA	10863
Qy	5703	TTGACATCTCTGATGACGCTTTTGTGATCATCAGATCATCAACATTTTGAAGTTA	5762
Db	10864	TTGCATCTCCGAAACGCTGCTTTTATCAGGACATCAAGATCACCACCTGCTCAACAGTCA	10922
Qy	5763	GCCTGCACAGTACGACACTGACTTATTTCTGCAACACTTGGTGTGTTCTAACTTACAT	5822
Db	10924	AAATTGATGTCAAGTGTGATGACTTATTTACGCGAATTCGAGGAGATGCTAACCTTCAGT	10983
Qy	5823	ACAAAGCTGACAGGAGGAGGACATTTGTCTCAAGTTTCACTCCACCTTCAAGACAGAGCTGTTTGA	5882
Db	10984	ATGATATCCGACCCGGAAGGACATATGCGCTGTATATTCGACTTGCAGACACAGCAACCTCC	11043
Qy	5883	AGGAAGGACACACATGTGACTCTCCGTATGAGCAGCATTAACACTATCTTTAGCATTCGA	5942
Db	11044	AAGATGTGACAGTTCATGTCTCTGGAAGAAAGAACCGGTGACAGTACCTTCAGCACCGCGA	11103
Qy	5943	GCCCAACAGCAATTTTATATGTTTTCGTATGCGGCAAGAAAGTCCACTGCATATGCTGAAT	6002
Db	11104	GCCCAACAGGAACTTCATGTATTCGTGTGTGTAAGAAAGCAACATGCATGACGAAT	11163
Qy	6003	GTAACCAACCGCGCGACACATATATTTGAGAGAACACATAAAGTCGACCAAGAAATTCAGG	6062
Db	11164	GCAAAACCAACAGCTGATCATATGCTGAGACCCCGCACAAAAATGACCAAGAAATTCAGG	11222
Qy	6063	CGGCGATTTCCAAACACTTTGGAACCTGCTGTTGCATCTGTTGGGGAGCATATCCC	6122
Db	11224	CCGCGATCTCAAAAACCTTCATGAGATGTGGGTGTGTTCCCTTTTCGCGCGCGCTCTGCTCC	11283
Qy	6123	TCATTTGTTTGAAGACTTATATGTTGTGTCTGCAAGCTTATGCTTATAAACACACGTGAAT	6182
Db	11284	TATTAATTAATGAGACTTATATATTTTGTTCGACAGATGATGCTGATACACACGAAGAT	11343
Qy	6183	GACTGACCGCGGACACTGA-CATATGCGGTAAACCTGATGACTTCCGAGGAAGCTGTGT	6241
Db	11344	GACCGCTAACGCCCAATGACCCGACAGCAAAACTGATGTACTTCCGAGGAACGTATGT	11403
Qy	6242	GCATTAATGCCACGCGCGCT-----TGACACT	6268
Db	11404	GCATTAATGCATACAGGCTGTATATTAAGATCCCGCTTACCGCGGCAATATAGCACACC	11463
Qy	6269	AAAACTGATGTATTTTCCAGGAAGCACATGATATATGTGTGCAGTGT	6318

Db 11464 AAACTCGAGTATTTCCGAGGAAGCGAGTCATTAATGCTGGCAGTGT 11513

RESULT 4

US-09-367-764-7

Sequence 7, Application US/09367764

Patent No. 6583121

GENERAL INFORMATION:

APPLICANT: Johnston, Robert E.

APPLICANT: Davis, Nancy L.

APPLICANT: Simpson, Dennis A.

TITLE OF INVENTION: System for the In Vivo Delivery and

TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Bell Seltzer Park & Gibson, P.A.

STREET: 1211 East Morehead Street

CITY: Charlotte

STATE: No. 6583121th Carolina

COUNTRY: USA

ZIP: 28234

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/367,764

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/801,263

FILING DATE: 19-FEB-1997

ATTORNEY/AGENT INFORMATION:

NAME: Sibley, Kenneth D.

REGISTRATION NUMBER: 31,665

REFERENCE/DOCKET NUMBER: 5470-147

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-420-2200

TELEFAX: 919-881-3175

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 1163 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-09-367-764-7

Query Match 37.0%; Score 2397.6; DB 4; Length 1163;

Best Local Similarity 64.6%; Pred. No. 0;

Matches 3752; Conservative 0; Mismatches 1984; Indels 74; Gaps 9;

Db 580 ACTGACGCTATGAAACGGGAGCGTATATTTCTATCGGAAACAGGCCAAGTCACTTC 639

5707 ACTGCTAACCGGGAGTAGGTGATATTTTCACGACAGAGCCCTGGGACCTTGC 5766

640 AACGAAATCACTAGTCAATGTAATCAAGAACTTATTTGATGGGCGCTCCATG 699

5767 AAAAATACTCGCTTTCAGAAACAGCTTACAGAAACCGACCTTGAAGGCAATGTTCTGG 5826

700 AGAATATTAGCGCCCGCGCTCGATCTCGAAGAGAAATGTTACGAAGAACTGC 759

5827 AAAGATCTACCGCCCGCGCTCGACAGTCAAGAAAGAGACAGCTCAAACTCAGGTACC 5886

760 AATTATGCGCTCTGAAGAAATAGAGAGGATCAATCAAGAAAGTAGAAATATGA 819

5887 AGATATGCGCCCGCGAGCAAAAGAGAGTACAGTCTGAAAGTAGAAAGAAACAGA 5946

820 AAGCAATTAACAGGAGGAGCTATTTCTGATTTGGGCAATATCTATCATGAGAAGTGA 879

5947 AAGCATTAACCACTAGGAGGAGCTTTTCAAGGAGCTAGGCTGTATTAATCTGCACAG--- 6003

Qy 880 ATCTGTGAGTGTATACAGAGTCAATTAATCTGTACCAATCTACTGTCAACGTAATTA 939

Db 6004 ATCAGCCGAATATGCTATTAAGATCACTTACCGAAACATGCTATTCAGAGCTGATCCAG 6063

Qy 940 ACAGTTTACATCTGCAAGAGTTCGGGTTAAACGTCAACTTAATTAATCAAGAAAT 999

Db 6064 CGAATCTCTGACCCAAAGTTTCTGTAGCTGTGTTTAACTATCTGATGAGAAAT 6123

Qy 1000 ACCCTAGAGTCCAGTATGTTATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 1059

Db 6124 ACCGACGTTAGCATCTTATCAATCAACAGAGTACAGATGCTTACTTGGATATGATG 6183

Qy 1060 ACGGCGATCTGTCTGTATGATACAGCCACTTTTTCGGCTTAAGTGAAGACCTACC 1119

Db 6184 ACGGACAGTGGCTTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 6243

Qy 1120 CAAGAGCATAGTATTTGACAGCCAGATTAAGATTAAGATTAAGATTAAGATTAAG 1179

Db 6244 CGAAAGACAGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 6303

Qy 1180 ATACATTAAGAAATGATTTGCTGTGAGTCACTTAAGAAATTAAGATTAAGATTAAG 1239

Db 6304 ACAGTTGCAAAAGCTGTCAATGCTGAGTCACTTAAGAAATTAAGATTAAGATTAAG 6363

Qy 1240 GAGAAATTAAGTCTGTATGATTCGGGCGATTTAATGTTGTTCAAGAAATTAAGAT 1299

Db 6364 GTGAATCTCCAACTGAGATCTGAGATCACTTAAGATTAAGATTAAGATTAAGATTAAG 6423

Qy 1300 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1359

Db 6424 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6483

Qy 1360 TTACCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1419

Db 6484 TTACCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6543

Qy 1420 ATAAATCTAAACCGTTGAGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 1479

Db 6544 ATAAATCTAAACCGTTGAGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 6603

Qy 1480 ATGTCAAGTATCTCCGAGCAAGAACTATCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1539

Db 6604 ACGTGAAGATTAACCTGAGCAAGAACTATCAAGAGAGAGAGAGAGAGAGAGAGAGAG 6663

Qy 1540 AGGCTGAGATCCCTGCTGATACCGCTTACCTTCCGAGATCAATGCGGAATTAAGTCCGTA 1599

Db 6664 AAGCGGAGAAACCTGCGAGACCGCTTACCTTCCGAGATCAATGCGGAATTAAGTCCGTA 6723

Qy 1600 GACTGAATGCGGTGCTTCTGCAAAATATCAATCTCTTGCACATGTCAGCGGAAGATT 1659

Db 6724 GCTTACAGCCGTTTGTGATACCAATCACTTCAAGCTTGTGATGATGATGATGATGATGAT 6783

Qy 1660 TTGATGCGATTAATGCTGAACATTTCCACAGCGGAGCCAGTATTTGGAACGAGCATGC 1719

Db 6784 TTGATGCGATTAATGCTGAACATTTCCACAGCGGAGCCAGTATTTGGAACGAGCATGC 6843

Qy 1720 CGTGTGATTAAGAAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1779

Db 6844 CGTGTGATTAAGAAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6903

Qy 1780 ACTTAGGATGAGCAACCGCTTGAATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1839

Db 6904 ACTTAGGATGAGCAACCGCTTGAATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6963

Qy 1840 CTGTGACCTTAACAG 1899

Db 6964 CCACCAATCTGCGCCAGGAGTCCGTTCAAAATTCGGGGAGATGATGAATCCGGAATGT 7023

Qy 1900 TCTTAAGGCTTTTGTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1959

Db 7024 TCTTAAGGCTTTTGTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7083

Qy 1960 AACGTTTAACAGCTGAGCGTGCAGGCTTATTCGAGAGATTAAGATTAAGATTAAGAT 2019

Db 7084 AGCGGCTTTAAACGTCCTCAAAATGTGACAGATTATTCGGGAGACGACAACTTTATACCGAG 7143
Qy 2020 TCGCTCCGACACCTTGATGCGGAGAGATCGCCACTTGGCTGAACATGGAAGTAAAA 2079
Db 7144 TAGATCTGACAAAGAAATGCTGAGAGTGGCCACTGGCTCAACATGGAAGGTATAGA 7203
Qy 2080 TTATGATGACGTTAATTTGATGATCAAGACACCTTCTGTGGGGAGATTTATCCGTGGG 2139
Db 7204 TCATTGACGAGCATCGGCGAGAGACCACTTACTTCTGCGGTGATTCATCTTGGCAG 7253
Qy 2140 ACCAGATPACAGGACAGCCTGACAGAGTGCAGACCCCTCTAAAAAGGCTTTTAAAGCTTG 2199
Db 7264 ATTGGTTACTCCACAGCGTGTGCGGTGGGAGCCCTTGAAAAGGTGTTTAAAGTTGG 7323
Qy 2200 GAAACACATTGCGCATGATATACCAAGCTGCGACCGCCCGGGCACTGCATGATG 2259
Db 7324 GTAAACCGCTCCAGCCGACATGACMAAGACGAAGACGAAGACGCGCTCTGCTAGATG 7383
Qy 2260 AAGCAATGCGATGGAACGAATTTGGAATTAAGGACGATTTAGTAAGGCGGTAGAAATCA 2319
Db 7384 AAACAAAGGCGTGTGTTAGATGATGATPACAGACCTTAGCAGTGGCCGTGGACATC 7443
Qy 2320 GATACGAGATCATCTGCGACGCGCTGATCATCGTCTGTGTCACAGTTAGCCGAAAGCG 2379
Db 7444 GGTATGAGGTAGACAAATCAACACTGTCTGTGGCATTTGAGAACTTTGCCAGAGCA 7503
Qy 2380 TTAAAGAACTTCAAGACATTAAGAGGAGCCCAATCACCCTTACGGCTGACCTTAATGG 2439
Db 7504 AAAGACATTTCAAGCCATCAGAGGGGAAATTAAGCATCTCTACGGTGTCTTAAATAGT 7553
Qy 2440 TGACGTAGTAGA-----CAGGACCTACCACCGGAGAAATGTTTCA 2482
Db 7564 CAGCATATGATCTTCACTCTACTAATACCAACACACACCACTGATGATGAGATTTCT 7623
Qy 2483 TACCCTCAGCTGAACCTTTCAACAGTTTACCTTCAATTCGATGCTTACCGAGATCCA 2542
Db 7624 TTAAATGCTGTGGCGCGCCCTTCCAGCCCACTGCATGTGAGGCGCGGAGAA 7683
Qy 2543 AACCTCTAGGCGCGCTGAGGCGCTTTGGGCGCCCGCTGGCTGCTCAATGGAAGAT 2602
Db 7684 GGAAGCAGGCGCGCCGATGCTCTCCGCAATGAGGCTTGCTTCCAAATTCAGCAACTGA 7743
Qy 2603 CTTAGAGGATGATAGTCACTTGCACTTTCAAAACAC--GATCACTTAATCGCGCGAG 2660
Db 7744 CCAAGCCGTTAGTCCCTTACTTGAACAGGCACTGACCTTAAACCCACGCGCCAC 7803
Qy 2661 GTCCACCGCAAAGAGAGAGAGTGTCTCTTAAAGCCAAAACCTTACTGAGCTTAAAAAG 2720
Db 7804 GCCCGCGCGCGGCAAGAGAGAGGCGGCCAAGCAACCCGAAAGCGAAAGCA 7863
Qy 2721 AGAAGCAGCAACCAAGAGAGCAAAAC--GCMAAGCTTAACAGGAAAGCAAGCACTA 2777
Db 7864 AAACACAGGAAAGAGAGAGAGCACTGCAAAACCCMAACCGGAAAGAGACAGCTTA 7923
Qy 2778 TGTGATGAAGTGGAGTGGAGCAAGACATTTCCGATC---ATGCTGAAGGCGCAAGTA 2834
Db 7924 TGGCATTTAAGTGGAGGCGGACAGACTGTTCAGAGTAAAAAATGAGAGCGAGATGCA 7983
Qy 2835 ATGATATGCTGCTGTGTCGAGAGAGCTGATGAACCACTCCACGTTGAGGAAAA 2894
Db 7984 TCGGGACGCACTGCGCATGAGAGAAAGATATGAACCACTCCACGTGAAGAGACTA 8043
Qy 2895 TTGATATGAGCAATTAGCGCGCGTGAATTTGAAGAGGCTAGCATGTCAGACTTTGGAGT 2954
Db 8044 TTGACACCCCTGTGTCTTCAAGCTCAAAATTCACCAAGTCTGACATACACATGTAGT 8103
Qy 2955 ACGGAGAGCTTCCCGAGAACTGAATGACAGGCTGACAGTACACCGGACAAACAC 3014
Db 8104 TCGCAGCTTCCCGGTCAACATGAGAGTGAAGGCTTACCTTACACGATGAACACCTTG 8163
Qy 3015 CGGGCTTCTAACATGCGACCAAGCGCGAGTCCAGTATGAGAAATGGAGATTTACCGTAC 3074

Db 8164 AAGGTTCTAACACTGACCAACGAGCGGTGACGATTAAGTAGAGGACGATTTACATCC 8223
Qy 3075 CGAAGGAGTGGCGGGAAGGAGACAGGGAAGACCGATCTGGAACAGAGGCGAG 3134
Db 8224 CCGCGGAGTAGAGGAGAGAGACAGTGTGTCTCGATTTATGATTAATCTCAGGCGGG 8283
Qy 3135 TTGTGCTAATTTGTTAGAGAGTGCAAATGAGGACAGCGCTTACGCGCTTTCAGTGTCA 3194
Db 8284 TTGTGCGATTAAGTCTCGAGAGGGCTGATGAGGAAACAAACCGCCCTTTCCGTGCTCA 8343
Qy 3195 CTGGAACCAAGAAAGGAGTACATTAAGGATACCCCGAAAGTTCTGAACCGTGT--- 3251
Db 8344 CTTGAATTAACAAAGGAGAAACATCAAGCAACCCCGGAAGGAGACGAAGAGTGTCTG 8403
Qy 3252 -----CACTAGTTACAGCGCTATGCGTGTTCGAAATGTCAGGTTCCATGGACAAAC 3305
Db 8404 CTGCACCACTGCTACCGGCAATGCTGTGTTGAAACGTAGCTTCCATGCAATGCC 8463
Qy 3306 CACCGTGTCTATTCACTGACCGCAAGAAACAACTCGACGTCTGGAAGAGACGTG 3365
Db 8464 CGCCCAATGCTACACCGCGGAACATCCAGAGCTCTGCATCTCGAAGAGACGTGA 8523
Qy 3366 ACAATCCAAATTACGACAGCTGTGAGAAAGCTTTGAAAT--GTCTATACCGCGGC 3422
Db 8524 ACCACGAGGCTTACGACACCTGCTCAACGCCATATTGCGGTGCGGATCTCGGCGAGAA 8583
Qy 3423 CCAAGCAAGCATTCAGGATGACTTCACTGACGACGCTCTCTGCGGAGTTCTGCGCGT 3482
Db 8584 GTAAAGAAACGCTACGACGACTTTCCTTGACACGCGCTGACTTGGGACATGCTGT 8643
Qy 3483 ATTGCAACACTCAACCGCTGTGTTTCAAGCCCAATTAATAATTGAGAACGTGTGGACGAAT 3542
Db 8644 ACTGTACCACTATCGAACCGTGTGTTAGCCGATTAAGATCGAGCAGTCTGGGATGAG 8703
Qy 3543 CTGATGATGATGATTAAGATTCAGGTCTGCGCAATTCGGCTTACATTCAGCAGGCA 3602
Db 8704 CGGACGACCAACCATTCGCTACAGACTTCCCGCCGCTTGGATGAGACCAAGCCGAG 8763
Qy 3603 CTGCGAGTGCACCAATTCGTTTACATGTTCTTTCGACGACGACATGACATCAAGAG 3662
Db 8764 CAGCAAGCTCAATTAATTAATACCGCTACATGTGCTTCGACAGATCATCTCAAGAG 8823
Qy 3663 ACAATGAGAAATAATGCTATGACACATCTGGAACCTCGCTCTTGGCCACAAAG 3722
Db 8824 GCACATGATGATGATCAATGATGACACCTCAGAGCCGTGTAAGAGCTTAGTACAAAG 8883
Qy 3723 GGTATCTTCTGTAGCTCAATGTCCTCAGGTGACAGTGAACCTGATATCAAGCGC 3782
Db 8884 GATACTTTCTCTCGCAAGTGTCTTCAGGGGACAGCTTACCGTATGACATAGCGAGTA 8943
Qy 3783 GAGCATCTGAGAAATTCATGACCGTGAAGAAAGATCAGAGAGATTTGTGGTGAAG 3842
Db 8944 GCAACTGAGAAAGTCAATGACAAATGSCCGCAAGATPAAACCAAAATTTGTGGAGCGG 9003
Qy 3843 AGGAGTACTTGTTCACACCGCTCATGGAAGGTGTGAAGTGCACAGTTTACGATCACT 3902
Db 9004 AAAAATATGACATCACTCCGTTTACGTTAAGAAAGATTCCTTGACAGTGAAGACGCTC 9063
Qy 3903 TGAAGAGAGCTGTGCGGAGTACATAACATGACAGGCGCAACGCGCTATTAAGT 3962
Db 9064 TGAAGAAACAAACGCGCGGTACTACTATGACAGGCGGAGCCGCAAGCGCTATATCAT 9123
Qy 3963 CTTATCTGAGAGAGCGTCAAGGGAAGTGAATTAACCACTTCTGGGAGAGAGCTCA 4022
Db 9124 CTTATCTGAGAGATTCATCAGGGAAGTTTACGGAAGCAACATTCGGGAGAGAACATTA 9183
Qy 4023 CTTACGAATTTAAGTGTGGGACTACAGACAGGTATCTGAGACCGGACGAAAGATGA 4082
Db 9184 CTTACGAGTCAAGTGGGAGTTTACAGACCGGAAACGTTTACGACCCGTTACGAAATCA 9243
Qy 4083 ACGGCTGCAATTAAGCAAAACAGTCAATTCCTTACAAAGCGACCAAAAGAAATGGTCT 4142
Db 9244 CGGGCTGACCGGCATCAACAGAGTGCCTGCTATTAAGAGCGACCAAGAAAGTGGTCT 9303

QY 4143 TCAACTGCGCGGATCTTATTAAGGACACAGACCACTCAGTGCAGATTAATTCACATTC 4202
DB 9304 TCAACTGCGCGGATCTGATTCAGACACGCGACCAACGCGCCAAAGGAATTCATTC 9363
QY 4203 CATTCGCTGGAACCGACAGTCTGCGGTTCCGTTAGTTCACACGCTCAGTACGA 4262
DB 9364 CTTCACGCTGATCCCGAGTACCTGATGTCCTGTTGCGCACGCGCGAAAGTATAC 9423
QY 4263 AGTGTAAAGGATCAACCTCCACCTGACCTGCAATGCGAACCACTTTCGTAACAGA 4322
DB 9424 ACGGTTTAAACACATCAGCTCCCAATTAGACACAGACCATCTGACATTCCTCACACA 9483
QY 4323 GAAATTTGGGCTGCGACAGACAGCAAGAGATTAACAGGCTCAGTACAGAGA 4382
DB 9484 GGAAGCTAGGGGCAACCCGGAACCAACCTGATGATCATCGGAAACAGGTTAGAA 9543
QY 4383 ATTTTCTGTGGGGGAGAAAGGCTGAGTACGTAAGGGTAAACATGAACCAAGTCAGAG 4442
DB 9544 ACTTCACCGTGCACCGAGATGGCTGGAATACATATGGGGCAATCAGAACAGTAAAGG 9603
QY 4443 TCTGGGCGGAGAGTGGGACACAGGCGACCAATGATGGCGGATGATCATCATCC 4502
DB 9604 TCTATGCCCAAGAGTCTGACACGAGAACCTCAGAGATGGCAGACAGAAATAGTACAGC 9663
QY 4503 ACTATATCATCGGATCAGTCTACATGTCATTTGTCTGTGTGCTGCTCTTGCTA 4562
DB 9664 ATTACTATCATCGCATCTGTGTACACATCTTACCGCTGCGATCAGCTGTGTGGGA 9723
QY 4563 TCTGTAGGCACTGCATCATCAGAGCTTGATCGCCAAAGCAAGAAAGACTGCTGA 4622
DB 9724 TGATGATGGGTAACCTGTGGCATATATGCTGTAAAGCGCGCTGTAGTGTCTGA 9783
QY 4623 CGCATTAAGCCTTGCACCGAAGCGAAGGTAACCCAGCATTAAGCGTTTGTGTGCA 4682
DB 9784 CGCATTAAGCCTTGCACCGAAGCGAAGGTAACCCAGCATTAAGCGTTTGTGTG 9843
QY 4683 TTGCGCAACCAACGCTGAAACATTTGGAAACCTTGAACCATCTGTGTTTAAACAC 4742
DB 9844 TTAGTGTGGCTAATGCTGAACATTCACCGAGACCATAGATTAATTTGTGCAACACC 9903
QY 4743 AACGTTTCTGTGGGACAGTTGTGATCTCTGTGACGCGTGTGTAATTCGTCCCT 4802
DB 9904 AGCGTTTCTGTGGGTCAGCTGTATACCTCTGCGCGTGTGTCTGTTAATGGCT 9963
QY 4803 GCTTTTATGCTGATGCTTTTATTTATGTTGTCAGGGCTCTGCTGGGAAAGTATACG 4862
DB 9964 GTTGTATGCTGCTGCTTTTATTTATGTTGTCAGGGCTCTGCTGGGAAAGTATACG 10023
QY 4863 CTTTGAACATGCGACAGCTGTGCCAAATGTTCCGGGATCCCGTAATAAGGCTGTGCG 4922
DB 10024 CTTTGAACATGCGACAGCTGTGCCAAATGTTCCAAATGTTGCCAAGATACCGTAAGGACTTGTG 10083
QY 4923 AACGCGAGTTTACGCGGCACTTAACTGAGATCAGCGTCTCATCGAATTAACAC 4982
DB 10084 AAAAGGAGGGGTAAGCCCGCTCAATTTGAGATTAATGTCATGTCCTGGAGGTTTGC 10143
QY 4983 CTTCACTTAACAAGAGTATGTAAGCTTGAATTTCAACAGTCAATTCCTTACACAG 5042
DB 10144 CTTTCAACCAACCAAGATTAATCTGCAATTCACACATGTTGCTCCCTCCCTTAAAG 10203
QY 5043 TTAATGTGCGGCTCCTCGATGCAAGGATCTCTAAAGCGGATTAACATGCGCGG 5102
DB 10204 TCAATGTGCGGCTCCTTGAATGTCAAGCCGCTCTACGCAAGCTATACCTGCAAGG 10263
QY 5103 TTTTGGGAGTGTATACCTTTTATGTTGGGAGGCGCAATGCTTCTGTACAGTGA 5162
DB 10264 TCTTTGAGGGGTTATACCTTTATGTTGGGAGGCGCAATGTTTGGGAGTGA 10323
QY 5163 ACACACACTAGTATGAGGCTATGTCGAGTTGCTTCCGATCTGATTAATGATCAGAG 5222
DB 10324 ACAGCAGATGATGATGAGGCTATGTCGAAATTTGTAGTATGTCGCGAGCTACGACGCGC 10383

QY 5223 TCGCACTAAAGTTTACACAGCTGCTGTAAAGTGGCGCTGCTGATATAGTACGCAACA 5282
DB 10384 AGGGATTAAGTGCATATCTGCCGATGAAGATAGACTGCTGATATGTAAGGAAAC 10443
QY 5283 CCACCGGCGACCTGATATGTTTGTCAATGGCGTACGCGAGTTCTTACGGGACTGA 5342
DB 10444 CTACAGTTTCTTATGATGTATGTAAGAGTACACAGGAAAGCTTAAAGACTGA 10503
QY 5343 AGCTATGACAGGCGGATATACCGGCTTTTACCTTTGACCATTAAGTCCGATCA 5402
DB 10504 AAGTATAGCTGACCAATTTACATTTGTTTAAACATTTGATCAGAGGCTGTTATCA 10563
QY 5403 GAAAGGGCTTGTATTAACATACACTTCCCTGATATGAGTATGAACAGAGGCT 5462
DB 10564 ATCCGGCTGTGTATACATATGATCTTCCGAAATACGAGCATGAACAGAGGCT 10623
QY 5463 TCGGATATTAAGCATCTGTGTTGATGCTAACAGATATAGTACCGGATCATAC 5522
DB 10624 TTGAGACATTCAGATCTCTTGACTAGCAAGACCTCATGCGCGACAGACATTA 10683
QY 5523 GCGTGTGAAGCCTTGTGTCAAGAAATCCAGCTCCCTTACACCAAGCAGTATCAGG 5582
DB 10684 GGTACTTAAGCCTTCCGCAAGAACGTGATGTCGTTACACGAGGCGCATCTGAT 10743
QY 5583 ATGAATGTGGAAGAACTCAGAGACCCCTGCAAGAAACAGACCATTTGGATGA 5642
DB 10744 TCGAGATGTGGAAGAAACATCAGGCGCGCCATCAGAGAAACGCGCTTTGGTGA 10803
QY 5643 AAATTAAGTGAAGCTCTGCGAGCTTAACTGTGCTTACGCGGACATCTCTATCGA 5702
DB 10804 AGATTCAGTCAATCCCTTGCAGCGGAGCTGCTCATAGGGAACATTCCTCATTTCTA 10863
QY 5703 TTGACATCCCGAGCAGCTTTTGTAGATATAGATATCAACAAATTTTAAAGTTA 5762
DB 10864 TTGACATCCCGAGCAGCTTTTGTAGATATAGATATCAACAAATTTTAAAGTTA 10923
QY 5763 GCTCAGAGTACAGACTGATTAATCTGAGACTTGTGGTGTCTCTAATTAACAGT 5822
DB 10924 AATGTATGTATGATGATGATCTTATTCAGAGATATAGATATCAACCTGTCTCAAGTCA 10983
QY 5823 ACAAGCTGACAGGAGGAGATTTGTCAATGTCATCTCCACTCAGACAGCTGTTTGA 5882
DB 10984 ATGTATCGACCGGAGAGCAATGCTGTGATTCGATTCAGACACAAACCTTCC 11043
QY 5883 AGGAAGGACCAACATGTATCTGCGGAGCAGATTAACCTATTTTATGACATGA 5942
DB 11044 AAGGTGACAGTTCATGCTGTGAGAAAGAGGCGTGAAGTCACTTCAGACCGGGA 11103
QY 5943 GCCCACAAGCAATTTTATAGTTTGTGCTATGCGGCAAGAACTCAACCTGCAATGCTGAT 6002
DB 11104 GCCCACAAGCAATTTTATAGTTTGTGCTATGCTGTGTGTGTAAGAACATGCAATGCAAT 11163
QY 6003 GTAACCAACGCGCGACACATTAATGAGAACCAATTAAGTGCACCAAGATTTCCAG 6062
DB 11164 GCAAAACCAACGCTGATCATGTGTGAGACCCCGCAAAATAAGACCAAGATTTCCAG 11223
QY 6063 CGGAGTTTCCAAACATCTTGAACGTGCTGTGCACTGTTGGGGAGACATCATCC 6122
DB 11224 CCGCATCTCAAAAATTCATGAGGTGGCTGTTTGCCTTTTGGGGCGCCTGTGCG 11283
QY 6123 TCATTTGTGTGACTTATATGTTGTGCTGACCTCATCTTATTAACACAGATAGAT 6182
DB 11284 TATTTATTAATGAGCTTATATGTTTGTGCTGACATGATCTGATACAGAGAT 11343
QY 6183 GACTGACGCGGACACTGA-CATAGCGTAAACATGATGATCTTCCGAGGAGCGTGT 6241
DB 11344 GACCGTACGCGCCCAATGACCCGACAGCAAAATCTGATGTACTTCCGAGGAGTGTAT 11403
QY 6242 GCATATATCCACGCGCGCT-----TGACACT 6288
DB 11404 GCATATATGATCAGGCTGTATATTAATGCCCGCTTACCGCGGCAATATAGCAACAC 11463
QY 6288 AAACTGATATTTTCCGAGGAACAGTGAATATGCTGTGACGTGT 6318

Db 11464 AAAAGTGGATTTTCCGAGGAGGCGAGCTAATGCTGGCAGGT 11513

RESULT 5
US-08-446-932-1
Sequence 1, Application US/08446932
Patent No. 5639650
GENERAL INFORMATION:
APPLICANT: Johnston, Robert E.
APPLICANT: Simpson, Dennis
APPLICANT: Davis, Nancy L.
TITLE OF INVENTION: CDNA Clone for South African
TITLE OF INVENTION: Arbovirus No. 5639650 86
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenneth D. Sibley
STREET: Post Office Drawer 34009
CITY: Charlotte
STATE: NC
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,932
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-118
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 881-3140
TELEFAX: (919) 881-3175
TELEX: 575102
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11663 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
US-08-446-932-1

Query March 36.9%; Score 2396; DB 1; Length 11663;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3751; Conservative 0; Mismatches 1985; Indels 74; Gaps 9;
QY 580 ACTGACGGTATGAAGCGGAGCGTATATTTTCATCGAAGACAGGCCAAGCTCACTTC 639
Db 5707 ACTGTCAACCGGGGTAGTGGGTACATATTTTCGACGACACAGGCCCTGGGCACTTGC 5766
QY 640 AACGGAATCGTACGTCAATGTAACTCAAGAACCTATATTGGATGGGCCCTCATG 699
Db 5767 AAAAGAGTCCGTTCTGCAAGAACAGCTTACAGAACCGACCTTGGAGGCAATGTTCTGG 5826
QY 700 AGAAGTATTACGCCCGCGCTCGATCTCGAAGAGAGAAATGTTACAGAAAGAACTGC 759
Db 5827 AAAGATTTACGCCCGCGCTCGATCTCGAAGAGAGAAAGCACTCAAGTCAAGTACC 5886
QY 760 AATTATGCGCTCTGAAGGAATAGAAAGCAGTATCATCAAGAAAGTAGAAATATGA 819
Db 5887 AGATGATGCCACCGAGACCAAAAGCAGTACCACTCGAAAATGAAAAACAGA 5946
QY 820 AAGCAATTACGCGGAGCATCTTTTGGATTGGGCAATATCTATCATCAGAAATGA 879
Db 5947 AAGCATTAACCATGAGCGACGTCTTTCAGGGCTACGACTATTAACCTCTCCACAG--- 6003

QY 880 ATCTGTGAGTGTACAGAGTCATATTCCTGTACCAATCTACTGTCAACCGTAAATTA 939
Db 6004 ATCAGCAGAAATCTCTAATTAAGATCACTACCCGAAACCATGTATTCAGAGGTACAG 6063
QY 940 ACAAGTTTACATTCGAGAGGTGCGGTTAAACGTGCAACTTATGTTATCAAGAGATT 999
Db 6064 CGAAGTACTGTGACCCAAAGTTGCTGTAGCTGTTTGTAACTATCTGCAAGAGATT 6123
QY 1000 ACCCTACAGTACGAGTATTTATTAACAGATGAATACGATGCGTATCTTGACATGGTGG 1059
Db 6124 ACCGACGATGAGATCTTATCAAGATCACGACGATGACATGCTTACTTGATGTGTAG 6183
QY 1060 ACCGCGATGCTGTCTGTAGATACAGCACTTTTTCGAGTAACTGAGAAGCTACC 1119
Db 6184 ACCGACAGTGGCTTGGCTGTAGATACGCACTTTTTCGAGTAACTGAGAAGTACC 6243
QY 1120 CAAAAGACATACATTTTTCAGCCAGATTAAGATCAGCCGTCCTATACAGA 1179
Db 6244 CGAAAGACACGAGTATAGAGCCCAACATCCGAGTGGGTTCCATCAGCATGCGAGA 6303
QY 1180 ATACATTACAAATGATTTGGCTGACGCTACTAAAGAAATGCAAGTACCAATGCG 1239
Db 6304 ACAGTTGCAAAACGTGCTATTGCGGCACTAAAGAAATCTCAAGCTACACAAATGC 6363
QY 1240 GAGAAATTAAGTCTTAGATTTCGCGGCACTTATATGTTGATTTTCAGAAATACGAT 1299
Db 6364 GTGAATCCGACACATGAGTCAAGCGCATTAACGTTGAATGCTTTGCAAAATATGAT 6423
QY 1300 GCAATGATGATCTGGGATACCTTTCGATTAACCTTATGCGCTACTACAGAGAACG 1359
Db 6424 GCAATGACGATATGGGAGAGTTTCCGAAAGCAATTAAGATCACTACGAGTTCCG 6483
QY 1360 TTAACGAATATGTGACAAAGCTGAAAGGCGGAAAGCAGAGCATTTGTCGAAATCTC 1419
Db 6484 TTACCGCATCGTGGCCACATTAAGAGCCCTTAAGCCGCGCATGTTGCAAAAGACGC 6543
QY 1420 ATATCTMAAACCGGTGACAGAGATACCAATGATCAATTCGTATGATGATCTAAAGAG 1479
Db 6544 ATATTTGTTCCATTTGCAAGAGTGTCTATGATGATGATGATGATGATGATGATGATG 6603
QY 1480 ATGTCAAGTTACTCCGCGACGAAACATACAGAGAGCGGCTTAAGTGCAGTTATTC 1539
Db 6604 ACGTGAAGTTACACCTGGCAGCAACACACAGAAAGAAACCGAAAGTCAAGTGTATC 6663
QY 1540 AGGCTGAGATCCCTGCTGACCGCTTACCTTTGCGGAGTCAATCGGAAATTAAGTCCGTA 1599
Db 6664 AAGCGCAGAACCTTGGCAGCGCTTACTTATGCGGATTCACCGGAGATTAAGTGGCA 6723
QY 1600 GACTGAATGCGGTGCTTTCGCAAAATATCCATCTCTTTCGACATGTCAAGCGGAATTT 1659
Db 6724 GCGTTACAGCGTTTTCCTACCAACATTCACACGCTCTTTGACATGTCCGCGAGAGACT 6783
QY 1660 TTGATGAGTATTTGCTGAACATTTTCACACAGCGGACCCAGATTTGGAAACGGACATCG 1719
Db 6784 TTGATGCAATCAACAGCAACATCTTAAGCAAGGTGACCGGATCTGGAACGGATATTCG 6843
QY 1720 CGTGTGTAATAAAGGAGAGAGAGCTATGCGCATTTGGGCGTTGATGATCTTGAAG 1779
Db 6844 CCTGTTGACAAAGCAAGCAAGAGCGCTATGCGGTTAACCGGCTGATATCTTGAAG 6903
QY 1780 ACTTAGGTGACCAACCGCTCTTATGATTTGATAGAGGCGGCTTTCGCAATATCAAT 1839
Db 6904 ACTGAGGTGAGCAACCACTACTGACTGATGAGTGGCGCTTTGGAAATATCAT 6963
QY 1840 CTGTGACCTTACTACAGAAAGAGGTTTAAATTTGCTGCAATGATGAATTCGGATGT 1899
Db 6964 CAACCATCTGCGCCACCGGTATCCGTTTCAAAATTCGGGAGATGATGAATTCGGAAATGT 7023
QY 1900 TCTTAAGCTGTTTGTGAACAACATGCTCAATATCATTTGCTAGCAGAGTACTAGTG 1959
Db 7024 TCTTCAAGCTCTTGTGAACAAGTCTGATGATGCTTATGCGCAGAGATTTGAGAG 7083
QY 1960 AACGGTTAACAGGTCAAGCGTGGCGGCTCTTATCGCGCAGATTAACATAGTGCATGTG 2019

Db 7084 AGCGCTTAAACGTCCTCAAAATGTGCAAGCATTTATCGCGACGACGAACTTATACACGAG 7143
Qy 2020 TCGTCCGACACCTTGATGGGAGAGATGGCCACTTGGCTGAACATGAGTAAATA 2079
Db 7144 TAGATCTGACAAAGAAATGGCTGAGAGGTGTGCCACTGCGCTCAACATGAGGGTTAAGA 7203
Qy 2080 TTATGATGCAATTTATGGTATCAAGACCTTACTTCTGTGGGGATTTATCTGTGG 2139
Db 7204 TCATGAGCACTATCGCGAGAGACCACTTACTTCTGCGGTGATTTCACTTGGCAAG 7263
Qy 2140 ACCAGATTAACAGGCAACGCTGACAGTGCAGACCTCTTAAAGGCTTTTAAAGCTTG 2199
Db 7264 ATTCGGTAACTCCACAGCGGTGTCCGTGGCGGACCCCTTGAAGAGCTGTTAAGTTGG 7323
Qy 2200 GAAACCAATGCGCAATGATGATACCAAGATGGAGCCCGCGCGGACCTGCATGATG 2259
Db 7324 GTAAACCGCTCCACCGCAGATGAGCAAGACGAAAGACGAGCGCTGTGTAGATG 7383
Qy 2260 AAGCAATGCGATGAAACAGAAATTTGAAATTAACGAGATTAAGGCGGTAGAAATCA 2319
Db 7384 AAACAAAGCGGTGTTAGAGTATTAACAGACACTTACAGATGGCCGCGGCAATC 7443
Qy 2320 GATACGAGATCACTACTGCGAGGCTGATCATCACTCTCTGTCAAGTTAGCCGAAAGG 2379
Db 7444 GGTATGAGGTAGACAACATCAACCTGTCTGTGCAATGAGAACTTTTGGCCAGAGA 7503
Qy 2380 TTAAAGACTTCAAGAGCATTAAGAGGAGCCCAATACCTCTACGCGTGAACCTTAATAG 2439
Db 7504 AAAGGCAATTTCAAGCCATCAAGAGGGAATTAAGCATTTCAAGGTGTCTTAATAGT 7563
Qy 2440 TGACCTAGTAGA-----CAGCAGCTTACCCAGCGGAGAAATGTTTCCA 2482
Db 7564 CAGCATAGTACATTTCACTGTGACTATTAACCAACACCAACCACTGATAGAGATTTCT 7623
Qy 2483 TACCTCACTGAACCTTTCCACCACTTAACCTTACCAATCCGATGCTTACCGAGATCA 2542
Db 7624 TTAAATGCTCGGCGCGCGCCCTTCCAGCGCCCACTGATGTGAGAGCGCGGAGAA 7683
Qy 2543 AACCTCTAGGCGCGCGCTGAGAGCGCTTGGCGCCCGCTGGCTGCTCAATGCAAGAT 2602
Db 7684 GAGGACAGCGCGCCCGGATGCTGCGCCGCAATGAGGCTGGCTTCCAAATCCGCAACTGA 7743
Qy 2603 CTTAGAGAGTGAATGATCACTTGAATTTCAACAAC--GATCACCTAATCCGCGCCAG 2660
Db 7744 CCACAGCGGTCACTGCTTACTGATTTGAGACAGGCACTAGACTTCAAAACCCACCCAC 7803
Qy 2661 GTCCACCGCCAAAGAAAGAAAGATGCTCTTAAGCCAAACCTTACGAGCTTAAAGAA 2720
Db 7804 GCCCGCGCGCGCGCAAGAAAGCAAGGCGCCAAACCAACCGAAGCCGAAGAAACAA 7863
Qy 2721 AGAAGCAGCAAGCCAGAGAGCAAAAC--GCAAGCTTAACAGGAAACGACAACTTA 2777
Db 7864 AAACCAAGAGAAAGAAAGAAAGCACTGCAAAACCCAAACCCGAAAGAGAGCTTA 7923
Qy 2778 TGTGATGAAGTGGAGTCCGACAGACATTTCCGATC--ATGCTGAACGCGCCAGTGA 2834
Db 7924 TCGCATTAAGTTGAGGCGCAACAGATGTTGACGCTCAAAATAGAGACGAGATGCA 7983
Qy 2835 ATGATATGCTGCTGCTTTCGAGAGAGGCTGATGAACCACTCCACGTTGAGGAAATA 2894
Db 7984 TCGGACGACATGCGCATGAGAAAGTAAATGAACCACTCCACGTTGAAGAAACATA 8043
Qy 2895 TTGATATAGACAAATTAAGCGCGCTGAAATTTGAAGAGGCTAGCATGTACGACTTGAAGT 2954
Db 8044 TTGACACACCTGTGCTATCAAAATCTCAAAATTCACCAAGTGTCAACATACGACATGAGT 8103
Qy 2955 ACGGAGCGTTCCCGAAGCATGAATGACAGCTGAGTACACGCGCAACACAC 3014
Db 8104 TCGCACATGTCGGGTCAACATGAGAAAGTGAAGGCTTCACTTACACAGTGAACCTTG 8163
Qy 3015 CGGCGTTTACAACTGGACCAACGCGGACGTCAGTATGAGAAATGGAGATTTACCGTAC 3074

Db 8164 AAGGTTCTAACAATGACACCAAGGCGGTGACAGTATAGTGAAGGACAGATTTACATCC 8223
Qy 3075 CGAGAGAGTGGCGGGGAAAGCCGACACCGGAAAGCCGATCTCTGACAAACAGAGCAAG 3134
Db 8224 CCGCGGAGTGAAGGAGCAGAGAGACAGTGTGTCTCCATTAATGATTAATCAAGCGCGG 8283
Qy 3135 TTGTGCTAATTTGTTCTAGAGAGTGCAAATGAGAGGACCGCTACGCGCTTTCAAGTCA 3194
Db 8284 TTGTGCGCATAGTCTCTCGAGAGGGGTGATGAGGAAACAAAGACCGCTTTTCGTCGCA 8343
Qy 3195 CTGGAACCAAGAAAGGAGTCACTTAGAGATACCCCGAAAGTTCTGAACCGTGT-- 3251
Db 8344 CCTGGAATAGCAAAAGGAAAGCAATCAAGACAAACCCCGAAAGGACAGAAAGTGTGTG 8403
Qy 3252 -----CACTAGTTAAGGCGTATAGCGCTTGTGAATGTACGTTCCATGCGCAAAAC 3305
Db 8404 CTGACACCTGTCAGGCGCATAGTCTGTGTTGAAAGTGAAGCTTCCATGCAATCCGC 8463
Qy 3306 CACCGTGTGCTAATTTCACTGACGCGCAGAACGAACTGAGTGTCTGAAGAGAACTGCG 3365
Db 8464 CGCCCAATGCTTACACCCCGCAACCATTCAGAGCTCTGACATCTTCAAGAGAACTGA 8523
Qy 3366 ACAATCAATTAACGACACGCTGTGAGAAAGTCTTGAAT--GTCCATACGCGCGC 3422
Db 8524 ACCAGAGCGCTTACGACACCTGCTCAACGCGCATATTCGCTGCGGATGTCCGCGAGA 8583
Qy 3423 CCAACGAAAGATTAACGATATTCACATGACCACTGACCTTACTGTGGGCTTTCGCCGT 3482
Db 8584 GTAAAGAAAGGTCACTGACACATTTTACCTTGACAGCGCGTACTTGGGCACTGTCTGT 8643
Qy 3483 ATTCAGACACTCAACGCGGTGTTTACGCGCCCAATTAATTAAGAACTGTGGAGCAAT 3542
Db 8644 ACTGTCACTATGTAACCGGTCTTTAGCCGATTAAGATGAGACAGTCTTGGATGAG 8703
Qy 3543 CTGATGATGATGATTAAGATTCAGAGTCTGCGCAAAATTCGCTCAATCAGCGAGCA 3602
Db 8704 CGGACGACAAACACATAGCATACAGACTTCCGCGCAATTTGATACGACCAAGCGAG 8763
Qy 3603 CTGCGGATGATCAACAAATTCGCTTACATGTTCTTTCGACACGACATCAATCAAGAAAG 3662
Db 8764 CAGAAAGTCAATTAAGATGACGCTACATGCTGACGAGAGATCATCTGTCAAAAGAG 8823
Qy 3663 ACAGATGAGAAATAGCTTACGACACATCTGACCTGCGCGTGTCTTGGCCCAAG 3722
Db 8824 GCACTATGATGATCAATCAAGATCAGACCTCAGAGCGGTGATGAAGGCTTTGATCAAAAG 8883
Qy 3723 GGTACTTCTGTTAGCTCAATGTCTCTCAGGTGACAGTGAACGTCAGTATCAAGACGCG 3782
Db 8884 GATATCTTCTCTCGCAAGTGTCTCTCAGGGAACGTTAGCATGAGAGATA 8943
Qy 3783 GAGCATCTGAGATTTCACTGACCGGTGAGAAAGAAAGATGAGAGAGTGTGTGTTAGAG 3842
Db 8944 GCAACTGACAACTGATGACAAATGCGCCGCAAGATTAACCAAAATTCGTGGACCGG 9003
Qy 3843 AGAGTACTTGTTCACCGCGCTCATGGAAGCTGTAAAGTGAAGCAAGTTTACGATCACT 3902
Db 9004 AAAAATATGACCTTACCTCCGTTACGATGAAGAAATCTTTCGACACATGACACCTC 9063
Qy 3903 TGAAGAGACGTTCTGCGGGTACATTAACATGACACAGGCGCAGCGCCACCGCTAATAGT 3962
Db 9064 TGAAGAAACAAACCGCGGCTACATCACTATGACACAGCGCGGAGCCGATGCTATACAT 9123
Qy 3963 CTTATCTGAGAGAAAGCTGAGGCGAAGTATCAATTAACCACTTCTGCGCAAGAGCTCA 4022
Db 9124 CTTATCTGAGAGAAATCAATGAGGAAAGTTTACGGAAGCCACATCTCCGGAAGAAACATTA 9183
Qy 4023 CTTAGATGATGATGATGAGCTTACAGACAGGATGTGAGACGCGGAACGAAAGATGA 4082
Db 9184 CTTAGATGATGATGAGGCTTATTAAGACCGGAAACCTTATGACACCTGACCGAAATCA 9243
Qy 4083 ACGGCTGACCTTAAGCAAAACAGTGAATGCTTCAAGAGCGACCAACGAAATGGGTCT 4142
Db 9244 CGGCGTGAACCGCATCAAGAGTGTGCTGTATTAAGAGGACCAAAACGAAATGGGTCT 9303

QY 4143 TCACTGGCGGATCTTATTAGGCAACAGACCACTGAGTGAAGTAAATTGACATTC 4202
DB 9304 TCAACTGGCGGATCTCGATGAGACCGCGACACACCGCCCAAGGAATTTGATTC 9363
QY 4203 CATTCGCTTGACACCGACAGTCTCCCGGTTCCGTTAGTCAACGCTTACAGTCA 4262
DB 9364 CTTTCAAGCTGATCCCGAGTACCTGACATGTCCTGTTGCCCAAGCGCGAAGCTAGTAC 9423
QY 4263 AGTGGTTCAAGGATCACTCTCACTGACTGCAATGCGACCAATTTGTTGACAA 4322
DB 9424 ACGGCTTTAAACATCATAGCTCTCAATTAGACACAGACCAATCTGATCTGTCAC 9483
QY 4323 GAAATTGGGGCTCGAGACAGCAACAGCAAGATGATTAACGGGCTTACCTCAGGA 4382
DB 9484 GGAACATAAGGGGCAACCCGGAACCAACCTGATGATCATGGAACAGGGTTAGAA 9543
QY 4383 ATTTTCTGTGGGGCGAAGAGGCTGAGTACGTATGGGGTAAACCAATGACAGTCA 4442
DB 9544 ACTTCACTGCGACGAGATGGCTGGAAATCATATGGGGCAATCAAGAACATTAAGG 9603
QY 4443 TCTGGGCGGAGAGTCCGACACAGCCACCATGATGGCGCATGAGATCATCC 4502
DB 9604 TCTATGCCAAGAGTCTGACACAGAGACCTCAAGATGGCCACAGAAATAGTACAG 9663
QY 4503 ACTATATCATCGGATCCAGTCACTGTCATTTGCTGTGTGGTGGCTTGGCTA 4562
DB 9664 ATTACTATCATCGGATCTGTGTGACACATTTAGCCGTGCACTACGCTGTGGCGA 9723
QY 4563 TCCGTGAGGCACTGCATCATCAGACAGTTGTCATCGCCAAAGCAAGAGAGTCCGTA 4622
DB 9724 TGATGATTTGGGTAACGTGTTGACAGATTAATGTCCTGTAAGCGCGGTAGGCTGA 9783
QY 4623 CGCCATACGCGCTTGACACCAAGCAAGGATACCAAGATTAAGCGGTTTGTGCTGA 4682
DB 9784 CGCCATATGCGCTCGCCCAATGCGGATTTCCAACTTCGCTGCACTTTTGTGCTGT 9843
QY 4683 TTGGCCAAACCAAGGCTGAACATTTGGAGAACTTTGAACCATCTGTGGTTGAACA 4742
DB 9844 TTAGGTGGCTTAATGCTGAACATTTACCGAGACATGATTTACTTTATGTCGAACGC 9903
QY 4743 AACGTTCTCTGGGCAAGTTGTGATTCCTCTGAGAGCGCTTGTATTCTGTTCGCT 4802
DB 9904 AGCGTTCTTCTGGGTCAAGTGTGATTACTCTGGCGCTGTGTGTCTATAGGCT 9963
QY 4803 GCTTTTATGCTGATGCTTTTATTATTGTTGCAAGCGCTGTGCTGGGAGGTAAG 4862
DB 9964 GTTGTCTATGCTGCTGCTTTTATTAGTGTGTCGGCGCTTACCTGGCGAAGGTAAG 10023
QY 4863 CTTTGAACATGGACACAGTGTGCAATGTTCCGGGATCCCGTATTAAGGCGTTGGTC 4922
DB 10024 CTTAGAACATGGACACAGTGTGCAATGTTCCGGGATCCCGTATTAAGGCGTTGGTC 10083
QY 4923 AACGCGAGGTTAGCGGCACTTAACCTGAGATCAAGGTCGTCTACCGAATTAAC 4982
DB 10084 AAAGGCGAGGTTAGCGGCTCAATTGAGATTAATGTCATGCTCCGAGGTTTTGC 10143
QY 4983 CTTTAACTAACAAGAGTACGTGACCTGCAATTTCCACACAGTATTCCTTACCAAG 5042
DB 10144 CTTTCAACCAACCAAGAGTACGTGACCTGCAATTTCCACACAGTATTCCTTACCAAG 10203
QY 5043 TTAATGCTGGGGTCCCTGAGTGAAGGATCTCAAGAGGCGGATTAACATGCGCGC 5102
DB 10204 TCAATGCTGGGGTCCCTGAGTGAAGTCAAGCCCGCTCAAGCAAGATTAATCTGCAAG 10263
QY 5103 TTTTGGCGGTGTACCTTTTCAATGTTGGGAGGCGCAATGCTTCTGTGACAGTGA 5162
DB 10264 TCTTTGAGGGGTATACCTCTTCAATGTTGGGAGGAGCAATGTTTTTGCGACGTGGA 10323
QY 5163 ACACACACTGAGTGAAGCGTACGTGAGTTCGCTCAAGACTGCACTATAGATACGAG 5222
DB 10324 ACAGCCAGATGATGAGGCGTACGTCAATTTGATGATTTGCGCAGCACTGACACGCG 10383

QY 5223 TCGCAATAAAGTTACACAGAGTCTGTGAAGTCCGCTGCGTATGATTAACGCAACA 5282
DB 10384 AGCGGATTAAGTGTGACATCTGCCGATGAAGATGAGATCTGTATGTTACCGCAACA 10443
QY 5283 CCACCGGCACTGTGATACGTTTGTCAATGAGCGCTCACGCCAGGTTCTCACGGGACTGA 5342
DB 10444 CTACAGATTTCTTAGATGTGATGAGTGAACGAGTCAACAGGAACTGTAAAGACTGA 10503
QY 5343 AGGTCAATGAGAGGCGGATTAACGCGCTTTTACCCCTTTGACCAATGAAGTGTCTCA 5402
DB 10504 AAGTCATGCTGAACCAATTTGAGCATTTGTACACCATTCGATCAAGAGTGTATCA 10563
QY 5403 GAAAGGGCTGTGTTTACCACTAGCACTTCCCTGATGATGAGGATGAACCAAGAGCGT 5462
DB 10564 ATGCGGCTGTGTGATCACTATGATCTTCCGAAATGCGAGCATGAACAGAGCGT 10623
QY 5463 TCGGCAATTAACGATCCCTGCTGATGCTACAGCATATGATGCCGCACTGACATAC 5522
DB 10624 TTGGAGATTTCAAGCTTACCTCTTGTGACTAGCAAAAGACTCATGCGCAGACAGCATTA 10683
QY 5523 GCGTGTGAAGCTCTGTGTAAGAAATCAAGTCCCTTACCCCAAGCATGATCAGGT 5582
DB 10684 GCGTACTCAAGCTTTCGCAAGAAAGTGTATGCTCCGTACAGCAGGCGCATTTGAT 10743
QY 5583 ATGAATGTGAAGAAACAATCTCAGACGACCCCTGCAAGAAAGACACATTTGATGTA 5642
DB 10744 TCGAGATGTGAAGAAACAATCTCAGGCGCCGCACTGCAAGAAACCGCCCTTTTGGTGA 10803
QY 5643 AAATGAATGAGAGCTCTGCGAGCGCTTATCTGTGCTTACGGGACATCCATCTGCA 5702
DB 10804 AGATTCAGTCAATTCGCTTTCGAGCGGTGAGTGTCTATACGGGAACATTTCCATTTCTA 10863
QY 5703 TTGACATCCCTGATGACGCTTTTGTGATCATCAGAAATCAACAATTTTGAAGTTA 5762
DB 10864 TTGACATCCGAAAGCTGCTTTTATGACAGCATCAGATGACCACTGTCTCAACATCA 10923
QY 5763 GCTGACAGTGAAGACTGCAATTTTCTGACACTTGTGTGTTCTTAACATTAAGT 5822
DB 10924 AATGTATGTCAAGTGAAGTGAATTTTCAAGCGGACTTCCGAGGATGCTTACCTGAGT 10983
QY 5823 ACAAGCTGACAGGAGGAGCATTTGTCAGTTCACTCCCATCCACAGCAAGCTGTTTGA 5882
DB 10984 ATGTATCCAGCCGGAAGGAGCAATGCGCTGTACATTTGCAATTCGAGACAGCAACCTCC 11043
QY 5883 AGGAAGGACACACATGATCTGCGGTAGGAGACATTAACATTAATTTAGACATGCA 5942
DB 11044 AAGAGTGAAGATTCATGCTGTGAGAAAGAGCGGTGACATTAAGACACGCGGA 11103
QY 5943 GCCCAAGCAAAATTTATAGTTTGTGCTATGCGGCAAGAGTCACTGCAATGTGAAT 6002
DB 11104 GCCCAAGGGAATCTTCAATTTGTATCGTGTGTGTGAAGAGACATGCAATGCAAT 11163
QY 6003 GTAAACCAAGCGGCGACACATTAATGAGAACCAATAAATGTCACCAAGAAATTCAGG 6062
DB 11164 GCAACCAACAGCTGATCATATGCTGAGCACCCCGCAAAATGACCAAGAAATTCAGG 11223
QY 6063 CGGCAAGTTCCAAACATCTTGAACATGCTGCTGCACTGTTTGGGGAGCATATCCC 6122
DB 11224 CCGCATCTCAAAAATCTTATGAGAGTGTGCTGTTGCTTTTGGGGGCGCTGCTGCG 11283
QY 6123 TCAATGTTGAGGACTTAATGATGTTGTGCTGAGCTTATGCTTATTAACACAGTGAAT 6182
DB 11284 TATTAATTAAGACATTAATGATTTTGTGCTGAGCATGATGCTGACAGCAAGAT 11343
QY 6183 GACTGACGCGGACACTGA -CATAGCGTAAACTGATGTACTTCCGAGAAAGCGTGT 6241
DB 11344 GACCGCTAACCCCAATGACCCGACAGCAAACTGATGTACTTCCGAGAACTGATGT 11403
QY 6242 GCATATGCAAGCGCGCT-----TGACACT 6268
DB 11404 GCATTAATGATCAAGGCTGTGATTAATGATCCCGTTACCGGCGGCAATTAAGCAACACC 11463
QY 6269 AAAACTGATGTATTTCCGAGGAAGCACAGTGATTAATGCTGTGAGTGT 6318

Db 11464 AAAAAGTCAAGTATTTCCAGAGAGGCGAGTGCATTAATGCTCGAGTGT 11513

US-08-801-263A-1
Sequence 1, Application US/08801263A

GENERAL INFORMATION:
PATENT NO. 5811407
APPLICANT: Johnstone, Robert E.
APPLICANT: Davis, Nancy L.
TITLE OF INVENTION: System for the In Vivo Delivery and
TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
STREET: 1211 East Morehead Street
CITY: Charlotte
STATE: No. 5811407th Carolina
COUNTRY: USA
ZIP: 28234

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,263A
FILING DATE: 19-FEB-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 5470-147
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11663 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 60..7559
NAME/KEY: CDS
LOCATION: 7608..11342
US-08-801-263A-1

Query Match 36.9%; Score 2396; DB 1; Length 11663;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3751; Conservative 0; Mismatches 1985; Indels 74; Gaps 9;

QY 580 ACTGACGATTAAGCGGAGCGCTATATTTCTCATCGGAAACAGGCCAGGTACCTTC 639
Db 5707 ACTGCTTAACCGGGGTAGGTGGTACATATTTTCAGCGACAGGCCCTGGGCACTTGC 5766
QY 640 AACAGAAATCAGTACGTCAATGTAACAGAAACCTATATTGATGGCGCGTCAATG 699
Db 5767 AAAAGAAATCGCTTTCGCAAGACCAAGCTTACAGAACCACTTGAAGCCCAATGTTTCG 5826
QY 700 AGAAGTATTAGCGCCCGCGCTCGATCTTCGAAAGAGAGAAATGTTTACAGAAATTCG 759
Db 5827 AAAAGATCTACGCCCGGTGTCGACAGCTGAAAGAGAGAAAGCTCAAACTCAGATACC 5886
QY 760 AATTATGCGCTCTGAAGAAATAGAGAGATTCATCAGAAAGTAAAGAAATATGA 819
Db 5897 AGATATGCGCCACGAAAGCAAAAGCAGGTACCAAGTCTGAAAGAAATGAAGAAACAGA 5946

QY 820 AAGCAATTACAGCGGAGGAGCTATTCTGATTTGGGACATATCTATCATCAGAGTGA 879
Db 5947 AAGCATATACCACTGAGGAGCTGTTTCAAGGGCTACGATGATATCTGCGACAG--- 6003
QY 880 ATCTGTGAGTGTATACAGAGTCAATTAATCTGTACCAATCTACTCTGTCACGGTATTA 939
Db 6004 ATCAGCCAGAAATGCTATTAAGATCACTTACCCGAAACATCGTATTCAGAGTATACAG 6063
QY 940 ACAGGTTTACATCTGACAGGTCGGGTTAAAGTCGAACTTAATTCAGAGAAAT 999
Db 6064 CCAATCTCTGACCCCAAGTTTGTGTAGCTGTTTGTAAACAATCTGACATGGAAT 6123
QY 1000 ACCCTACAGTACCGAGTATTTGATATACAGATGATGATGATGATGATGATGATG 1059
Db 6124 ACCGAGGTTGCACTTATATAGATCAGACGAGATGATGATGATGATGATGATGATG 6183
QY 1060 ACGGCGCATCGTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1119
Db 6184 ACGGAGCAGTGTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 6243
QY 1120 CAAGAAGCATAGCTATTTGACAGCAGATTAAGATCAGCGCTGCTGCTGCTGCTGCT 1179
Db 6244 CAAAAGACAGAGATAGAGCCCAAAATCCGAGTGCAGTCTCATCAGAGATGAGA 6303
QY 1180 ATACATTAACAAATGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1239
Db 6304 ACAGCTTCAAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6363
QY 1240 GAGAAATTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1299
Db 6364 GTGAAGTCCCAACATGAGTACGAGCAGCATTCAGTGAATGCTTGAAGAAATATGAT 6423
QY 1300 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1359
Db 6424 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6483
QY 1360 TTACGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1419
Db 6484 TTACGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6543
QY 1420 ATATCTTAAACCGTTGCAAGAGATACCAATGATCAATTTGCTCATGATCTTAAAGAG 1479
Db 6544 ATATCTTAAACCGTTGCAAGAGATACCAATGATCAATTTGCTCATGATCTTAAAGAG 6603
QY 1480 ATGCAAGTATCTCCCGGACAGAAACATACAGAGAGAGCGCTTAAGTGCAGTTATTC 1539
Db 6604 AGTGAAGATTAACATCTGCGACGAAACACAGAAAGAAACCGAAAGTACAAAGTATAC 6663
QY 1540 AGGCTGAGATCCCTGCTACCGGCTTACCTTTGGGGATCCATCGGGAATTAATGCGGA 1599
Db 6664 AAGCGGAGAAACCTTGGCGACCGCTTACCTATGCGGATCCACCGGAGTTAGTGCGA 6723
QY 1600 GACTGAATGCGGTGCTTTCGCAATATCAATCTCTTTCGACATGTCAGCGGAAGAT 1659
Db 6724 GACTGAATGCGGTGCTTTCGCAATATCAATCTCTTTCGACATGTCAGCGGAAGAT 6783
QY 1660 TTGATGCAATATGCTTAAACATTTTCACACGCGGACCGCATTTGGAAGCGACATCG 1719
Db 6784 TTGATGCAATATGCTTAAACATTTTCACACGCGGACCGCATTTGGAAGCGGATTCG 6843
QY 1720 CGTGTGTTGATTAAGAGCAAGAGCGATTCGCAATTTGGCGTGTGATGATCTTTGAGG 1779
Db 6844 CCGTGTGTTGATTAAGAGCAAGAGCGATTCGCAATTTGGCGTGTGATGATCTTTGAGG 6903
QY 1780 ACTTAGGTGACCAACCGCTTTAGATTTGATAGAGCGCGCTTTCGCAATATCACAT 1839
Db 6904 ACTTAGGTGACCAACCGCTTTAGATTTGATAGAGCGCGCTTTCGCAATATCACAT 6963
QY 1840 CTGTCACCTACCTACAGAGAGAGGTTAAATTTGTCGATATGAAATTCGGTATAGT 1899
Db 6964 CCACCATCTTCCACAGGCTTCCGTTTCAATTTCCGAGGATGATGAAATTCGGGAAATGT 7023
QY 1900 TCTTAACGCTGTTGTCAACACACTGATATATCATGATTTGTAGCAGAGTACTACGCTG 1959

```
Db 7024 TCCGACGCTCTTGTCAACAGCTGCGAATGTCGTTATGCGCAGCAGAGTAATTGGAGG 7083
Qy 1960 AACGGTTAACCAAGCTCAGCGCGCGGCTCTATCGCGGAGAGTAACATAGTGAATGTG 2019
Db 7084 AGCGGCTTTAAACCTCCAAAGTGCAGCATTTATTCGGGAGCAGAACATTTATACCGAG 7143
Qy 2020 TCGCTCCGACACCTTGTGCGGAGAGATGCGGCACCTTGTGCTGAACATGGAATTAATAA 2079
Db 7144 TAGATCTGACAAAGAAATGCTGAGAGTGTCCACTGCTGCTCAACATGAGGTTAAGA 7203
Qy 2080 TTATGTGACGATTATTGTATCAAGACCCCTACTTGTGCGGAGATTATCTGTGTG 2139
Db 7204 TCATTTACGAGCATCGCGCAGAGACCACTTACTTCTGCGGTGATTCATCTTTGCAAG 7263
Qy 2140 ACCAGATTAACGAGCAGACCTGCGAGAGTGCAGACCTCTTAAAAAGGCTTTTAAAGCTT 2199
Db 7264 ATTGCGTTACCTCCACAGCGGTGTGCGTGGCGAACCCCTTGAAAAAGCTGTTTAAAGTTGG 7323
Qy 2200 GAAAAACATTGCGCAGTGAATGCCAGACTGCGACCGCGCGCGGCACTGCAATG 2259
Db 7324 GTAAACCGCTCCAGCGCAGCATGAGCAAGACGAAAGCAGAAAGCGGCTTGTCTAGATG 7383
Qy 2260 AAGCAATGCGATGGAACGAAATTGAAATTACGAGCGAGTTAGTGAAGCCCTAGAAATCCA 2319
Db 7384 AAACAAAGGGGTGTTAGAGTAGTATTAACAGACACTTATGAGTGGCGGTGCAATC 7443
Qy 2320 GATTCGAGATCATCTGCGAGGCTGATCATCACTGTCTGTCCAAGTTAGCGGAAGCG 2379
Db 7444 GGTATGAGTAGACAAACATCACACTGTCTGTGGAATTGAAACTTTTCCCAAGACA 7503
Qy 2380 TTAAAGACTTCAAGAGCATTAAGAGGAGGCCAATCACCTCTAGCGGTGACCTTAATAG 2439
Db 7504 AAGAGCATTTCAAGCCATCAGAGGGGAAATTAAGCATCTTACGGTGTCTTAAATAGT 7563
Qy 2440 TGACGTAGTGA-----CAGCGACCTTACCCACCGGCGAAGTCTTCCA 2482
Db 7564 CAGCATGTACATTTCACTGACTAATACACAAACACACCATGATGATGAGATTTCT 7623
Qy 2483 TACCTTACGTGAATCTTCCACCAAGTTTACCTTAACAAATCCGATGCTTACCGGATCCA 2542
Db 7624 TTAAATGCTCGCGCGCGCGCTTCCAGGCCCACTGCATGTGAGGCGCGGAGAA 7683
Qy 2543 AACCTCTAGGCGCGCTGAGGCGGTTTGGCGCGCGCTGCTCAATGGAAGAT 2602
Db 7684 GAGAGCAGGCGGCGCGGCTGCTGCGCAATGGGCTGTGCTTCCCAATCCAGCACTGA 7743
Qy 2603 CTTAGAGGTGATGATCACTTGAACCTTCAACAAAC--GATCACTTAATCCGCGCAG 2660
Db 7744 CCAAGCCGTCAGTGCCTAGTCAATTGGAAGGCAAGCACTGAACCCACGCGCAC 7803
Qy 2661 GTCCACCGCCAAAGAAAGAAAGAGTGTCTTAAGCCAAACCTACTCAAGCTTAAAAAGA 2720
Db 7804 GCCCGCGCGCGCGCAAGAAAGAGGCGCCAAAGCAACCCAAAGCCGAAAGAAACAA 7863
Qy 2721 AGAAGCAGCAGCAGCAGGAGCAAAAC--GCAAGCCTAAACAGGGAAACGACAGTA 2777
Db 7864 AAACACGAGAAAGAAAGAAAGAGCACTGCAAAACCCGAAAGAGAGAGCGTA 7923
Qy 2778 TGTGTATGAAGTTGAGTGCAGCAAGACATTTCCGATC--ATGCTGAACGCGCAAGTGA 2834
Db 7924 TGGCATTAAAGTTGAGGCGCAGACAGCTGTTCAGCTCAAAATAAGAGAGAGATGTCA 7983
Qy 2835 ATGATATTCCTGCTGTGTGCGAGGAAAGGCTGATGAAACCACTCAAGTTGAAGAAAA 2894
Db 7984 TCGGCGCAGCACTGCGCATGGAAGAAAGATTAAGAAACCACTCAAGTGAAGAACTA 8043
Qy 2895 TTGATTAAGACATTTAGCGGCGGTGAATTGAAGAAAGGTAGATGAAGACTTTGAGT 2954
Db 8044 TTGACCACTCTGTCTTCAAAAGCTCAAAATTCACCAAGTGTCTGACATACAGCATGAGT 8103
Qy 2955 ACGGCGAGTTCGCCAGAAACATGAATCAGACGCTGCACTACCAAGCAGCAAAACAC 3014
Db 8104 TCGCACAGTTCCGGCTCAACATGAGAAAGTGAAGGCGTTCACTTACACCAAGTGAACCCCTG 8163
Qy 3015 CGGCTTTTCAACTGGCACACAGGCGAGTCCAGATATGAAATGGAGATTTAACGTTAC 3074
Db 8164 AAGGTTCTAACACTGGCACACGAGCGGTGCAATGATGAGGCGAGATTTACCATCC 8223
Qy 3075 CGAGAGTGTGGCGGGAAGGCGACACGCGAAAGACCGATCTTGACCAACAGAGCGAG 3134
Db 8224 CCGCGGAGTAGAGAGGAGAGAGACAGTGTGTGCTGCAATTATGATTAATCTAGCGCGG 8283
Qy 3135 TTGTGCTATTGTTCTAGAGGTGCAATATAGGCGACGCTTACGGCGCTTCACTGTGCTCA 3194
Db 8284 TTGTGCGATATGCTTCGAGGCGGTGATGAGGAACAAACCGCCCTTTCGTCCTCA 8343
Qy 3195 CTGGAACCAAGAAAGGGGTGACCATTAAGGATACCCCGAAGGTTGAAACCGTGT--- 3251
Db 8344 CTTGGAATAGCAAGGGAACAAATCAAGCAACCCGGAAGGACAGAAAGTGTCTG 8403
Qy 3252 -----CACTAGTTACAGCGCTATGCGTCTTTCGATGTCAAGTCCATGCGACAAAC 3305
Db 8404 CTGACACACTGTGTCAGCGGCATGTGTCTGTGGAACGTGAGCTTCCATGCAATGCC 8463
Qy 3306 CACCGTGTCTATTCACTGACCCAGAACGAACACTGACGTGCTGAAAGAAAGCTCG 3365
Db 8464 CGCCCATGTCTACCCCGGAACCATCCAGACTCTGCAATCTCGAAGAGAACCTGA 8523
Qy 3366 ACAATCAATTTACAGACGCTGTGAGAAAGCTTTGAAAT---GTCCATCACGCGCGC 3422
Db 8524 ACCACAGGCTTACGAAACCTGCTCAAGCCATATGCGGTGCGATGCTCCGCGGAA 8583
Qy 3423 CCAACGAAACATTAACGATGATGATCACTGACAGTCCCTTACTGCGGCTTGCCTG 3482
Db 8584 GTAAAGAAACGTGTCAGTGAAGCTTTAATCTTGACAGCCGCTGTTGGGCAATGCTGT 8643
Qy 3483 ATTGCAAGACTAAAGCGCGTGTTCAGCCCAATTAATAATTGAGAACGTGGGACGAAT 3542
Db 8644 ACTGTCACTATGTAACCGTGTCTTACCGCATTAAGATCGAGAGTGTGGATGAAG 8703
Qy 3543 CTGATATGATGATGATTAAGATCAGGTCTGGGCAAAATTCGCTCAATCAGGACAGCA 3602
Db 8704 CGAGCGCAACACCATGCGATACAGACTTCCGCGCAATTTGATGATGACCAAGCGAG 8763
Qy 3603 CTGCGATGTCAACCAATTCGTTAATGATCTTTTGAACCAAGCAATGATCAAGAGAG 3662
Db 8764 CAGAGCTCAATTAAGTACCGTCACTATGCTGTGAGAGATCAATGCTCAAGAGAG 8823
Qy 3663 ACAATATGAGAAATATGCTATGACACATCTGGAACCTGCGCTGTCTTGGCCAAAG 3722
Db 8824 GCACATGATGATCATCAATGATGACACCTCAGGACCGTGTAGAAAGCTTACCAAG 8883
Qy 3723 GGTACTTCCGTAGTCAATGTCTCCAGGTGACAGTGTAAACCGTCAATCAGGAGCG 3782
Db 8884 GATACCTTCTCTCGGAGATGTCTCCAGGGGACAGCGTAAGGTTAGATAGCGAGTA 8943
Qy 3783 GACATCTGGAATTAATGACACCGTGAAGAAAGATCAGAGAGAAATTTGTGCTAGAG 3842
Db 8944 GCAACTCAGCAAGTCAATGACATATGCCCCCGCAAGTAAACAAATTCGTGGAGCGG 9003
Qy 3843 AGAGATCTTGTTCACACCGCTCATGAAAGCTGTGTAAGTCCACGTTTACATCACT 3902
Db 9004 AAAAATATGACTTACCTCCGTTACAGGTAAAGAAATTCCTTGCACAGTGTACAGCCGTC 9063
Qy 3903 TGAAGAGAGTGTGCGCGGTATACATACATGACAGGCGCAGGCGCAACGCGTATAGT 3962
Db 9064 TGAAGAAACAAACCGCGGTATACATCTATGACAGGCGGGAACGCAATGCTTATACAT 9123
Qy 3963 CTTATCTGAGAGAGGTGAGCGAAGTATTAACCAACCTTCTGCAAGACGTC 4022
Db 9124 CTTATCTGAGAGATCATACAGGAAGTTTACGCGAAGCCACCATCCGGGAACAACTTA 9183
Qy 4023 CTTACGATTAATGTGTGCGACTACAGCAGATATCTGTAGCAGCGCAAGAGATGA 4082
Db 9184 CTTACGATTAATGTGTGCGACTACAGCAGATATCTGTAGCAGCGCAAGAGATGA 9243
```

4083 ACGGCTGCACTTAAAGCAAAACAGTCATTCCTTACAGAGCGAACCAACGAATGGGCT 4142
9244 CGGGCTGACCGGCATCAAGAGAGTGCGCTTATTAAGGCGACCAAGAGAGTGGCT 9303
4143 TCAATTCGCGGATCTTATTAGGCAACAGACCACTCAGTGAAGTAAATTGCACTTC 4202
9304 TCAACTTCGCGGATCTTATTAGGCAACAGACCACTCAGTGAAGTAAATTGCACTTC 9363
4203 CATTCGCTGACAGCGACAGTCGCGGCTTCGCTTACAGCGCTTACAGTACAGCA 4262
9364 CTTTAAGCTGATCCCGAGTACCTGATGCTCCGTTGCGCACGCGCAAGTATGAC 9423
4263 AGTGTTCAGAGGATCACTCCCTCACTGACATGAGCAACATTTGTCAGCAAGA 4322
9424 ACGGCTTAAACATCAAGCTCCCAATTAGACAGACCATATGACATTTGCTCACACCA 9483
4323 GAAATTTGGGCTGCGAGACAGCAAGCAAGAAATGATTAAGGGTCTACATCAGGA 4382
9484 GGAGACTAGGGGCAAAACCGGAACCACTGAATGATCATCGGAAACGCTTAGAA 9543
4383 ATTTTCGTGGGGGAGAGAGGCTGAGTACGTATGGGTATACATGAACAGTCAAG 4442
9544 ACTTACCGTGCACCGAGATGGCTGGAATACATATGGGCAATACAGAACCGATAGG 9603
4443 TCTGGGCGCAGAGATCGGACAGGCGACCCACATGATGGCGCATGATCATATCC 4502
9604 TCTATGCCAAGATCTGACACAGAGACCTTACAGATGGCCACACGAATATGACAGC 9663
4503 ACTATTCATCGGATCGATCTTACACTGTCATTTGCTGTGTGTGCTCTGTGTA 4562
9664 ATTTATTCATCGGATCGATCTTACACTGTCATTTGCTGTGTGTGCTCTGTGTA 4622
4563 TCTGTGATGAGCTGATCATACAGAGCTGATGCGCAAGAGAGAGTCTGCTGA 4682
9724 TGATGATGAGCTGATCATACGATGATGCTGATGAGAGAGAGTCTGCTGA 9783
4683 CGCCATGAGCGCTTGCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4682
9784 CGCCATGAGCGCTTGCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9843
4683 TTGCGGCAACCAAGCTGAAACATTTGAGAACTTTGAAACCATCTGTTGTTTAAAC 4742
9844 TTGAGTGGCTATGCTGAAACATTTGAGAACTTTGAAACCATCTGTTGTTTAAAC 9903
4743 AACCTTTCTCTGGGACAGTGTGATCTCTGTCAGCGCTTGTATTTCTGTTCCGT 4802
9904 AGCGTCTCTCTGGGACAGTGTGATCTCTGTCAGCGCTTGTATTTCTGTTCCGT 9963
4803 GCTTTTCATGCTGACAGCTTTTATTTGTTGAGAGAGAGAGAGAGAGAGAGAGAG 4862
9964 GTTGTCTATGCTGCTGCTTTTATTTGTTGAGAGAGAGAGAGAGAGAGAGAGAG 10023
4863 CTTTTCAGATGACAGCACTGTGCAAAATGTTTCGAGAGATCCCTGATTAAGGCTGTG 4922
10024 CTTTTCAGATGACAGCACTGTGCAAAATGTTTCGAGAGATCCCTGATTAAGGCTGTG 10083
4923 AACGCGAGTTTACCGGCACTTAACTTGAAGATCAAGGTCTCTCATTCGAATTAAC 4982
10084 AAAGGCGAGGTTACCGGCACTTAACTTGAAGATCAAGGTCTCTCATTCGAATTAAC 5042
4983 CTTTTCAGATGACAGCACTGTGCAAAATGTTTCGAGAGATCCCTGATTAAGGCTGTG 5102
10144 CTTTTCAGATGACAGCACTGTGCAAAATGTTTCGAGAGATCCCTGATTAAGGCTGTG 10203
5043 TTAATGCTGCGGCTCCTGAGTGAAGAGATCTCTCAAGGCGGATTAACATGCGCG 5102
10204 TCAGATGCTGCGGCTCCTGAGTGAAGAGATCTCTCAAGGCGGATTAACATGCGCG 10263
5103 TTTTGGGCGGCTGATCTTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5162
10264 TCTTTGAGAGGCTGATCTTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 10323

5163 AACACACATGATGAGGCGTACGTGAGTTGCTCCAGAGTCACTATAGTACAGGAG 5222
10324 AACGCAATATGATGAGGCGTACGTGAGTTGCTCCAGAGTCACTATAGTACAGGAG 10383
5223 TCGCATTAAGTTTCAACAGAGTGTCTGAAAGTGGCTGCTGATATATGCGCAACA 5282
10384 AGGCAATTAAGTGTATCTGCGCGAGTGAAGATGAGCTGATATGATGATGAGGAA 10443
5283 CCACCGGCGAGTATGAGTTGTCATGAGGTCAGAGGTTCCACAGGAGCTGA 5342
10444 CTACCATGTTCTTATGATGTGATCGTGAAGAGTACACAGGAACTGTAAGAGCTGA 10503
5343 AGTTCATGACAGGCGGATATCAGCGGCTTTTCAACCTTTGACCATTAAGTGTGATCA 5402
10504 AAGTATGATGAGGCAATTTACAGATGTTTATACCATTTGATATCAAGAGTGTATCA 10563
5403 GAAAGGCGCTGTTTACATCACTGAGCTTCTGATGATGAGCTATGAAACAGAGGCT 5462
10564 ATCCGCGCTGCTGATACATATGATCTTCCGGAATAGGAGCGATGAAACAGAGGCT 10623
5463 TCGGATATTCAGGATCCCTGCTGATGCTATGATGATGATGATGATGATGATGATG 5522
10624 TTGAGATATTCAGGATCCCTGCTGATGCTATGATGATGATGATGATGATGATGATG 10683
5523 GGTGCTGAGGCTTCTGTCAGAAACATCAAGTCCCTTACACCAAGCAGATATCAGGCT 5582
10684 GGTATCTCAACCTTCCGCGCAAGAGCTGATGCTCCGTACACGAGGCGGATTTGAT 10743
5583 ATGAATGTGGAAGAACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5642
10744 TCGAGATGTGGAAGAACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 10803
5643 AAATTTGAGGAGGCTTCTGAGAGGCTTATCTGCTTACAGGAGAGAGAGAGAGAGAG 5702
10804 AGATTTGAGGAGGCTTCTGAGAGGCTTATCTGCTTACAGGAGAGAGAGAGAGAGAG 10863
5703 TTGACATCTCTGATGAGAGCTTTTGTAGATATGATGATGATGATGATGATGATGATG 5762
10864 TTGACATCTCTGATGAGAGCTTTTGTAGATATGATGATGATGATGATGATGATGATG 10923
5763 GCTGACAGTACAGAGCTGATTTATTTCTGAGAGCTTTGTGTTCTTAACTTACAGT 5822
10924 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 10983
5823 ACAAGCTGACAG 5882
10984 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 11043
5883 AGAAGGAG 5942
11044 AAGAGTGAAG 11103
5943 GCCCAAGCAAAATTTTATGATTTGCTATGCGGAGAGAGAGAGAGAGAGAGAGAGAG 6002
11104 GCCCAAGCAAAATTTTATGATTTGCTATGCGGAGAGAGAGAGAGAGAGAGAGAGAG 11163
6003 GTAAACCAAGGCGGAG 6062
11164 GCAAAACCAAGGCGGAG 11223
6063 CGGAGATTTCAAAACATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6122
11224 CGGAGATTTCAAAACATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11283
6123 TCAATGTTGAGAGATTAATGATGTTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6182
11284 TATTAATTAATGAGAGATTAATGATGTTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 11343
6183 GACTGAGGCGGAG 6241
11344 GACGCTACAGGCGGAG 11403
6242 GCATATATGACAGGCGGCT-----TGACACT 6268

|||||
Db 11404 GCATTAATGCAATCAGCTGTGATATTAGATCCCGCTTACCGGGCAATATAGCAACACC 11463
|||
Qy 6269 AAAAAGTGTATATTTCCGAGGAAGCAGCTGCTATCTGTCTGCTGT 6318
|||||
Db 11464 AAAAAGTGTATATTTCCGAGGAAGCAGCTGCTATATGCTGCTGCTGT 11513
|||||
RESULT 7
US-09-102-248-1
Sequence 1, Application US/09102248
Patent No. 6008035
GENERAL INFORMATION:
APPLICANT: Johnston, Robert E.
APPLICANT: Davis, Nancy L.
APPLICANT: Simpson, Dennis A.
TITLE OF INVENTION: System for the In Vivo Delivery and
TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
STREET: 1211 East Morehead Street
CITY: Charlotte
STATE: No. 6008035th Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/102,248
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/801,263
FILING DATE: 19-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-147
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11663 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 60..7559
FEATURE:
NAME/KEY: CDS
LOCATION: 7608..11342
US-09-102-248-1
Query March 36.9%; Score 2396; DB 3; Length 11663;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3751; Conservative 0; Mismatches 1985; Indels 74; Gaps 9;

|||||
Db 5827 AAAAATCTAAGCGCCCGGTCTGACACGTCGAAAGGAAAGAGCTCAAACTCAGGTACC 5886
|||||
Qy 760 AATTATGCGCTCTGAAGAAATAGAAAGCAGGTATCAATACGAAAGATGAAATATGA 819
|||||
Db 5887 AGATGATGCCACCGAAGCCAAAGCAAGAGGATACAGTCTCAAAAAGTGAAGAACCGA 5946
|||||
Qy 820 AAGCAATTAAGGAGGAGCAGTCTATTTCTGATTTGGCAGATATCTATCATCAGAAAGTGA 879
|||||
Db 5947 AAGCCATTAACCATGAGCAGCTGCTTTTACGGGCTAGCAGCTGATTAATCTGCCACAG--- 6003
|||||
Qy 880 ATCTGTGAGTGTACAGAGTCAATATCTGTACCAATCTACTGTCACGCTAATTA 939
|||||
Db 6004 ATGACGCAAGTGTATTAAGATCACTTACCGAAACCATGATTTCCAGAGGTATACAG 6063
|||||
Qy 940 ACGGTTTACATCTGACAGAGTGGCGGTTAAACGTGCAACTTATGTTATCCAAAGAAAT 999
|||||
Db 6064 CGAATCTCTGACCCCAAGTTTGTGTACTGTCTGTTGTAACTATCTGATGAGAAAT 6123
|||||
Qy 1000 ACCCTACAGTACGAGTATTTATTAACAGATGAAATACGATGGATCTTGACATGGTGG 1059
|||||
Db 6124 ACCGACGAGTACATCTTATCAATACCGACAGTACGATGCTTACTTGGATATGTTAG 6183
|||||
Qy 1060 ACGGCGATGCTGTCTAGATACAGCCACTTTTGTCCGGCTAACTGAGAACTACC 1119
|||||
Db 6184 ACGGAGCAGTGGCTGTGCTAGATACGCACTTTTGGCCCCGCAAGCTTAGAACTTACC 6243
|||||
Qy 1120 CAAAGAAAGTACGATATTTTCCAGCAGAGTAAAGTACGCCCTCCCTATACAGA 1179
|||||
Db 6244 CGAAAGACAGAGTATAGAGCCCAAACTCCGAGTGGGTTCATCAGCAATGAGA 6303
|||||
Qy 1180 ATACATTAACAAATGATTTGGCTGAGCTACTAAAGAAATGCAAGCTTACCCAAATGC 1239
|||||
Db 6304 ACACGTTGCAAAAGCTGCTATGCTCCGAGCTAAAGAACTGCAAGCTCACCAAAATGC 6363
|||||
Qy 1240 GAGATTAACCTGTCTTATGATTCGCGGCAATTTATGTTGTTCAAGAAATACGCAT 1299
|||||
Db 6364 GTGAACGCCAACACTGAGCTGACGCAATTTCAACGTTGAATGCTTTGCAAAATATGCAT 6423
|||||
Qy 1300 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1359
|||||
Db 6424 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6483
|||||
Qy 1360 TTAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1419
|||||
Db 6484 TTAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6543
|||||
Qy 1420 ATAAATCTAAACCGTTGACAGAGATACCAATGATCAATTTGATGATGATGATGATGAT 1479
|||||
Db 6544 ATAAATTTGTTCCATTTGCAAGAGTGGCTATGATGATGATGATGATGATGATGATGAT 6603
|||||
Qy 1480 ATGTCAAAGTTATCTCCGCGACGAAACATACAGAGAGCGGCTTAAGTGCAGATTTC 1539
|||||
Db 6604 ACGTGAAGTTTACCTGCGACGAAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6663
|||||
Qy 1540 AAGCTGAGATCCCTGCTGCTACCGCTTACCTTTGCGGAGTCCATCGGGAATTAAGTCGTA 1599
|||||
Db 6664 AAGCTGAGATCCCTGCGACCGCTTACCTTATGCGGAGTCCATCGGGAATTAAGTGC 6723
|||||
Qy 1600 GACTGAATGCGGTGCTTCTGCAAAATATCCATCTCTTTCGACATGACGCGAAGATT 1659
|||||
Db 6724 GCGTTACAGCGGTTTCTGCAAAATATCCATCTCTTTCGACATGACGCGAAGATT 6783
|||||
Qy 1660 TTGATGCAATTAATGCTGAACATTTTCCACAGCGGACCAAGTATTTGAAGAGCACTCG 1719
|||||
Db 6784 TTGATGCAATTAATGCTGAACATTTTCCACAGCGGACCAAGTATTTGAAGAGCACTCG 6843
|||||
Qy 1720 CGTGTGTTAATAAGGAG 1779
|||||
Db 6844 CTTGTTGCAAAAGCAAG 6903
|||||
Qy 1780 ACTTAGTGTGACCAACCGCTCTTGAATTTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1839
|||||

Dh 6904 ACCTGGGTGTGAGCCAAACCACTATCTGACTTGATGAGTGGCCCTTTGGAGAAATATCAT 6963
Qy 1840 CTGTACACCTTACCTACAGAGAAAGAGTTAAATTTGTCATGATGAAATCCGGATGT 1899
Db 6964 CCACCATCTGTGCCACGGGTACCCGTTTCAATTTGGGGCCATGATGAAATCCGAAATGT 7023
Qy 1900 TCTTAAACGCTGTTTGTCAACACACTATGATCAATATCATGATTTGACAGAGTACTACGTG 1959
Db 7024 TCTTCAACCTCTTTTGTCAACACAGTTCTGAAATGTGTTATCTGACAGAGATTTGAGAG 7083
Qy 1960 AACGGTTAAACACGTGACAGGTGCGGGCTCTATGGGCGACATTAACATATGATGATGTG 2019
Db 7084 AGCGGCTTAAACGTCCTAAATGTGACATTTATCGGCGACCAACATTTATACCGAG 7143
Qy 2020 TCGTCTCCGACACCTTGTATGCGGAGAGATGCGCCACTTGAGTGAACATGAAAGTAAAA 2079
Db 7144 TAGTATCTGACAAAGAAATGCTGAGAGGTGTGCACTTGGCTTCAACATGAGGTAAAGA 7203
Qy 2080 TTATTTGATGACATTATGTTATCAAAAGCACCTTACTTCTGTGCGGAGATTTATCTGTGTG 2139
Db 7204 TCAATGACGACATCTGCGCGAGAGACCACTTACTTCTGCGGTGATTCATCTTGCAAG 7263
Qy 2140 ACCAATTAACAGGACAGCTTCAAGAGTGTGACAGCCCTTAAAGAGCTTTTAAAGCTTG 2199
Db 7264 ATTGGTTAACCCTCCACAGGCTGTGCGGTGCGAGCCCTTGAAGAGCTGTTAAAGTTGG 7323
Qy 2200 GAAACCAATGTCGATGATGATACCAAGACTGCGACCGCGCGGCACTGATGATG 2259
Db 7324 GTAAACCGCTCCACGCGACATGAGACAGAGACAAAGACGAGCGCTCTGTCTAGATG 7383
Qy 2260 AAGCAATGCGATGAAACAGAAATTTGAAATTAACGAGAGTGTAGTGAAGCGGTAGATCCA 2319
Db 7384 AAACAAAGGCGGTGTTAGAGTATTAACAGACACCTTACAGTGGCGGTGCACTGC 7443
Qy 2320 GATACGATCATATCTGCGAGGCTGATCATAGCTCTGTCTCAAGTTAGCCGAAAGCG 2379
Db 7444 GGTATGAGTATGACAAACATCACACCTGTCTGTGCTGCAATTTGAACTTTTGGCCAGAGA 7503
Qy 2380 TTAAGACCTTCAAGACATTAAGAGGAGCGCAATCACTCTAGAGGCTGACCTAAATAGG 2439
Db 7504 AAAGGCACTTTTACAGCCATCAGAGGAGAAATTAAGCATCTCTACGCTGTCTTAAATGT 7563
Qy 2440 TGACGTAGTAA-----CACGCACTTACCCAGCGGAGATGTTTCCA 2482
Db 7564 CAGCATAGTACATTTTCACTGTGACTAATACCAACACACACACATGATAGAGATCT 7623
Qy 2483 TACCTCAGCTGAATTTTCCACCAAGTTTACCTTAACATTCGATGGCTTACCGAGATCCA 2542
Db 7624 TTAACATGCTCGGCGCGCGCCCTTCCAGCGCCCACTGTCATGTGAGAGCGCGGAGAA 7683
Qy 2543 AACCTCTAGGCGCGGTGAGAGCGTTTGGCGCCCGCTGGCTGTCAATTCGAAGAT 2602
Db 7684 GGAGGACAGGCGCGCGAGTCTGTGCGCGCAATGGCTGGCTTCCAAATTCGACAACTGA 7743
Qy 2603 CTTAGAGGTGATATGATCACTTGAATTTCAAAAC--GATCACTAATCCGCGCGAG 2660
Db 7744 CCACAGCGGTGATGCTCTTACATTTGACAGGCACTAGACTTAAACCCACGCCAC 7803
Qy 2661 GTCCACCGCCAAAGAGAGAGAGTGTCTTAAGCCAAACCTTACGCTTAAAGAA 2720
Db 7804 GCCCGCCCGCGCGAGAGAGAGGCGCCAAACCAACCGAGAGCGAGAAACCA 7863
Qy 2721 AGAAGCAGCAAGCCAGAGAGAGAAAC---GCAAGCTTAAACAGAGAGAAACGACAT 2777
Db 7864 AAACAGAGAGAGAGAGAGAGAACTGCAAAACCCAAACCCGAGAGAGACGCGTA 7923
Qy 2778 TGTGTATGAAAGTTGAGTGTGAGAGAGATTTCCGATC---ATGCTGAACGCGCAAGTGA 2834
Db 7924 TGGCACTTAAGTTGAGGCGAGACAGTGTTCGAGTCAAAAATAGAGAGCGAGATGTCA 7983
Qy 2835 ATGATATGCTGCTGTGTGTGAGAGAGGCTGATAAACCACTTCAAGTTGAAGAGAAAA 2894
Db 7984 TCGGCGACGCACTGTGCTATGAGAGAAAGTAAATGAACCACTCCACGTAAGAGAACTA 8043

Qy 2895 TTGATATAGCAATTTAGCGCGCGTAATTGAAGAGGCTAGCATGTACGACTTGSAGT 2954
Db 8044 TTGACACCCCTGTGTATCAAAAGCTCAAAATTCACAAATCTGACACATACGACATGAGT 8103
Qy 2955 ACGGCGAGCTTCCCAAGCAATGAATACAGACGCTGCACTACACAGCGCAAAACAC 3014
Db 8104 TCGCACAGTTCGCGGTCAACATGAGAGAGTGAAGGCTTACCTACACAGTGAACCCCTG 8163
Qy 3015 CGGCTTCTAACAATGAGACCAAGCGCGAGTCCAGTATGAGAAATGGAGATTTACCGTAC 3074
Db 8164 AAGGTTCTTAAACCTGAGACCAAGAGGCGGTGACATATATGAGAGGAGATTTTACATCC 8223
Qy 3075 CGAGAGAGTGGCGGAGAAAGCGGACAGCGGAAACCGATCCTGGAACAGAGAGAG 3134
Db 8224 CCCCGGAGTGAAGAGAGAGAGACATGTGTCTGCTCAATATGATTAATCTAGGCGCGG 8283
Qy 3135 TTGTGGCTATTGTTCTAGAGGTGCAATGAGGCGACCGCTTACGCGCTTTCAGTGTCA 3194
Db 8284 TTGTGCGATAGTCTCGAGAGGCGGTATGAGAGAAACAAAGACCGCTTTCGCTGTCA 8343
Qy 3195 CTTGGAACAGAAAGGAGTACATTAAGGATACCCCGAAGTTCTGAACCGTGT--- 3251
Db 8344 CCGGAATAGCAAAAGGAGAAACATCAAGACACCCCGAAGGAGCAGAAAGTGTGTG 8403
Qy 3252 -----CACTAGTTACAGCGGTATGCGGTCTTTCGAATGTCAAGTTCCATGCGCAAA 3305
Db 8404 CTGACCACTGTGTACAGGCGCATGTGTCTTGTGAAAGTGAAGCTTCCATGCAATCGCC 8463
Qy 3306 CACCGTGTCTATTACTGACGCGAGACGAAACGACCTGACGTGTGCAAGAGAAAGTGTG 3365
Db 8464 CGCCACATGTGTACACCGGCAACCATCCAGAGCTCTGACATCTCTGAGAGAAAGTGTGA 8523
Qy 3366 ACAATCCAAATTAAGACACGCTGTGTGAGAAAGTGTGAAT---GTCAATACCGCGGC 3422
Db 8524 ACCAGAGGCTTACGACACCTGTCTCAACGCGCATATGCGGTGCGATGTGTCCGCAAA 8583
Qy 3423 CCAAGCAAGATTAACGATGATTTCACTGACGAGTCCCTGAGGAGTTTCCGCGT 3482
Db 8584 GTAAAGAGAGCTCACTGACATCTTACCTTACAGCGCGCTTACGAGTGTGTGAGAGTGTGT 8643
Qy 3483 ATTGCAACACTCAACGCGGTGTTACGCGCAATTAATTAATGAGACGTGTGAGAGAT 3542
Db 8644 ACTGTACCATATGAAACGATGCTTATGCGGATTAATGAGAGAGTGTGAGAGTGAAG 8703
Qy 3543 CTGATGATGATTCGATTAATTCAGAGTCTGTGCAATTTGGCTTCAATCAGCAGGCA 3602
Db 8704 CGACAGACACCAATACGATACAGACTTCCGCGCCAGTTTGAATACGACCAAAAGCGAG 8763
Qy 3603 CTGGGATGTACCAAAATTCGTTATCATGTCTTTCGACAGACGATGACATCAAGAG 3662
Db 8764 CAGCAAGCTCAATTAAGTACCGCTACATGTGCTGACAGAGATATATCTGCAAAAGAG 8823
Qy 3663 ACAGATGAGAGAAATAGCTATACAGACATGTGACCTGTGCGCTGTCTTGGCCAAAG 3722
Db 8824 GCAACATGATGATCAATGAATACAGACCTGAGACCGGTGAGAAAGCTTATGACTCAAG 8883
Qy 3723 GGTACTTCTGTAGCTCAATGTCTTCCAGGTGACAGTGAACGTCAGTATCAGAGCG 3782
Db 8884 GATCTTCTCTCGCGAAGTGTCTTCCAGGGGACAGGTAAACGTTAGCATAGCAGATA 8943
Qy 3783 GAGCATCTGAATTAATGACCGGTGAGAAAGAAAGTATGAGAGAAAGTTTGGTTAG 3842
Db 8944 GCAACTGACCACTGATGACCAATGAGCGCGCAAGATTAACCAAAATTCGTGGACGG 9003
Qy 3843 AGAGTACTTGTCCACCGCTCATGAGAAAGCTGTAAAGTGCACCTTTACATCACT 3902
Db 9004 AAAATATAGACTTACCTCCGTTACCGTAAAGAAATTCCTTGAACATGTATACAGCTTC 9063
Qy 3903 TGAAGAGAGCTTGTCCGGGTATCAATACATGACAGGCGCAGGCGCACAGCGTAAAGT 3962
Db 9064 TGAAGAGAAACAAACGCGCGGTATCACTATGACAGGCGCGGACCGCATCTATACAT 9123

QY	3366	CTATCTGGAGGAGCCCTCAGGCGCAAGTGTACATTAAACACCTCTCGGACAAAGTCGA	4022
Dp	9124	CTTATCTGGAGGATCATCTCAGGGAAGTTTATCGGAAGCCACATCTCGGGAAAGACATT	9183
QY	4023	CCTACGAATGAATGTGGCGACTACAGCAACAGGTATCTGAGACAACGAAACGAATGCA	4082
Dp	9184	CGTACGAGTGAATGGGGCGATTACAAAGACCGGAACCGTTAGACCCCGTACCGAAATCA	9243
QY	4083	ACGGCTGCATTAAGCAAAACATGCAATTGGCTTACANAGCGACCAAGCAAAATGGGTCT	4142
Dp	9244	CGGGCTGCACCGCCATCAAGCAGTGCCTCCCTTTAAGACCGACCAACCAAGATGGGTCT	9303
QY	4143	TCACCTGCGCGGATCTTTATAGGACAACAACACTCAGTGCAGAAATTTCCACATTC	4202
Dp	9304	TCAACTGCGCGGATCTGATCAGACACGCCACACAGGCCCAAGGGAAATTTCAATTTGC	9363
QY	4203	CATTCCGCTTACACCGACAGTGTGCCCGGTTCGGTAGTCTACACGCTACAGTCAACGA	4262
Dp	9364	CTTTCAAGCTATATCCGAGTACTCTGCATATGTTCCCTGTGGCCACGCGCCGAAAGTATGAC	9423
QY	4263	AGTGGTTCAAAAGGCATCACTCTCCACCTGCATGCAATGCGAACCAATTTCTGACAAACA	4322
Dp	9424	ACGGCTTTAAACATCATCAGCCTCCAAATTAGACACAGACACATCTGACATTTCTCACACCA	9483
QY	4323	GAAATTTGGGGCTGCCGACAGACCGCAACAGCAAAATGGAATTTACAGGTTCAKATCCAGGA	4382
Dp	9484	GGAACTAAGGGGCAAACTCCGGAAACCAACATGATATGATTCATCGMAAACACGGTTAGAA	9543
QY	4383	ATTTTTCGTGGGGCGAAGAGGCTGGAGTACGTATGGGGTTAACATCAATGAACCAAGTCAGAG	4442
Dp	9544	ACTTCAACCGTGCACCGAGATGGCTTGGAATACATATGGGCAATACAGAACCAAGTAAAGG	9603
QY	4443	TTTGGGCCCCAGAGTCGGCACACAGGCGAACCCCACTGAGTGGCGCATGAGATCATCATCC	4502
Dp	9604	TCATATGCCCAAGAGTCTGCACACAGAGAACCTCCACGATGGCCACACAGAAATAGTACAGC	9663
QY	4503	ACTATTATCATTCGGCATTCAGTCTACACTGTGCATATGTCGTGTGTGTGTGTGCTCTTGCTA	4562
Dp	9664	ATTACTATCATTCGGCATTCGTGTGTACACCACTTAAAGCCGTGCATCAGCTGTGTGGCGA	9723
QY	4563	TCCATGATAGGACATGCATCATCAGCAGCTTGCAATCGCCAAAGCAAGAAAGACCTGCCTGA	4622
Dp	9724	TGATGATTTGGCGTAACTGTTGCACACATTATGTGCTGTAAAGCGCGCGTAGTGCTTGA	9783
QY	4623	CGCCATACGCGCTTGCACCGAACCGAAACGGTACCACAGCATTAAGCGGTTTTGTGTGCTGA	4682
Dp	9784	CGCATATATGCCCTGGCCCAAAATGCCGATATCCAACTTGGCTGGCACTTTTGTGTGCTGTG	9843
QY	4683	TTGGGCCAACCAACGCTGAAAATTGTGGAGAACTTTGAACCATCTGTGGTTTAAACAAC	4742
Dp	9844	TTAGTGTGGCTTAAATGCTGMAAATTCCACCGAGACCATGATGTTACTTATGTGTCAACAGCC	9903
QY	4743	AAAGGTTTCTGTGGGACAGTTGTGCAATTCCTCTGGAGAGGCTTGTATTTCTGTCCGCT	4802
Dp	9904	AGCGGTTCTTGTGGTTCAGCTGTGTAACTCTCTGGCGCGTGTGTGTCTTAAATGGCT	9963
QY	4803	GCTTTTATGCTGTGATGCTTTTATTTATGTGTGAGCGCGTCTGCTGGGAGAAAGTATGACG	4862
Dp	9964	GTTGCTATATGTGCTGCTGCTTTTATGTGTGTGCGGCGCTACCTGGCAAGTATGACG	10022
QY	4863	CCTTGGAACATGCGACCACTGTGCCAAATGTTCCGGGGAATCCGGTATAGGCGTTGGTGC	4922
Dp	10024	CCTACGAACAATGCGACCACTGTGTTCCAAATGTGCCACAGATACCGTATAGGCACTTGTGT	10083
QY	4923	AAACGGCAGTTAAGCGCGCATTTAACTGTGAATCAAGGTGCTGTCAATGGAAATTAACAC	4982
Dp	10084	AAAAGGCGAGGATAGCCCCCTCAATTGGAGATTATGTATATCTCTCGAGAGTTTGTG	10144
QY	4983	CTTCAACTTAAACAAGAGTACGTGACCTGCAAAATTCACACAGCATCTCTTCAACACAAG	5042
Dp	10144	CTTCCACCAACCAAGATACATTTACCTGTCAAAATTTACACCATGTGTGTCTCCCTCCCTTAAG	10202
QY	5043	TTTAAATGCTGGGGTCCCTCGATGTCAAGGCAATCTCAAAAGCGGATTTACATATGCCCG	5102

Db	10204	TCAGATGCTGGCTCTTGGAAATGTCAGCCGCCCTCAGCAGACTATACCTGCAAG	10265
QY	5103	TTTTTGGCGGTGTGATCCCTTTTCATGTGGGAGCGCAATGCTTCTGTGACGTGAGA	5162
Db	10264	TCCTTGGAGGGGTGTACCCCTTCATGTGGGAGGAGCAAAATGTTTTCCGACAGTGAGA	10322
QY	5163	ACACACAACGTAGTGAAGGCTACGTGAGTTCGCTCCACACTCCACTATATAGTACCGGAG	5222
Db	10324	ACACCAGAAATGATGAAGCGCTACGTGGAATGTGCAATGATTCGGGACTGACACAGCGCC	10383
QY	5223	TCGCACATAAAGTTTCAACAGCTGCTCTGAAATCGCGGCTGCCTATAGTATACGGCAAC	5282
Db	10384	AGCGGATTTAAGTGTCAATCTGCCCGCATGAAGTAAAGTCTGGCTATATGTATCGGGAAC	10444
QY	5283	CCACCGCGCACTGGATACGTTTGTCAATGGCGTCAAGCCAGTTTCTCACGGGACTTGA	5342
Db	10444	CTACCAATTTTCCTAGATGTGTACGTGAAGGAGTCAACACAGGAACGTCTAAAGACTTGA	10500
QY	5343	AGGTCAATAGCAGGGCGCATATACGCCGCTTTTCAACCTTTTGACCTAAAGTGTCTATCA	5402
Db	10504	AAGTCATAGCTGTGACCAATTTTCAGACTGTGTTTACACATTTGATCAAGAGTCTGTATCA	10563
QY	5403	GAAGGGGCTTGTTTAACTACAGACTCTCCCTGAGTATGAGCTATGAACCGAGAGCT	5462
Db	10564	ATCGCGGCTTGTTGTACACTTATGACTTTCCGGAATATCGAGCGATGAACCGAGAGCT	10622
QY	5463	TCGGCGATTTTCAAGCATCTCTGCTTGAATGCTACAGACATATGATGCCGCACTGACATAC	5522
Db	10624	TTGGAGACATTTCAAGTACTCTCTTGATGCAAAAGACCTCATCGCAGACACAGACATTA	10683
QY	5523	GGCTGCTGAAGCTTCTGTCAAAACATCCAGTCCCCCTTACACCCAGCAGTATCAGGCT	5582
Db	10684	GGCTACTCAAGCCTTCGCGCAAAAGAGTGTATCCCGTACACAGCGCGCATCTCGAT	10743
QY	5583	ATGAATGTGGAAAGAACATCCAGAGCAAGCCCTGCAAGAAACAGACCACTTTGATGAT	5642
Db	10744	TCGAGATGTGAAAAACAATCCAGCGCGCCACTGACGAAACCGCCCCCTTTTGGGTGCA	10800
QY	5643	AAATTTGAAGTGAAGCCTCTGCGAGCGCTTAACTGTGCTTACCGGCAATCCTTATCTGCA	5702
Db	10804	AGATTGCAATGTCATCCGCTTCGAGCGGTGAGCTGCTATACGGGAACATTTCCATTTCTTA	10865
QY	5703	TTGACATCCCTGATGACGCTTTTGTGATCATCAGAAATCACCAACAATTTTGAAGTTTA	5762
Db	10864	TTGCATCTCCGAAACGCTGCTTATACAGGACATCAATCACCACCTGCTCTCAACAGTCA	10923
QY	5763	GCCTCCACAGTACGACATGATATTATTTGACACATTTGGTGTCTCTTAACTTACAT	5822
Db	10924	AATGTGATGTGATGATGATGCACTTATTCAGCGGACTTTCGAGGAGATGCTTACCCTGCACT	10983
QY	5823	ACAAAGCTGACAGGGAGGACATTTGTCTCAGTTTACTTCCACTCCACGACGCTGTTTTGA	5882
Db	10984	ATGTATCTCGAACCGGAAGACATATGCCCTGTATATTTGCAATTCGACACAGCAACCCCTCC	11043
QY	5883	AGGAAGGACCAACATGTGACTGCTCCGTAGCGAGCATTAACCTACTATTTAGCAATGCA	5942
Db	11044	AAGAGTGCAGATTTCAATGCTCCTGGAAGAAAGAACGGGTGACAGTACACTTCAGACCGCGA	11103
QY	5943	GCCCAACAAGCAATTTTATAGTTTTCGCTATAGCGGCAAGAAATCCACTCTGCATAGCTGAAT	6002
Db	11104	GCCCAACAGGCACTTCATTTGATTCGCTGTGTGAAGAAACAAATCAATGCATGCAAT	11163
QY	6003	GTAACACACGGCGCGACCATATTAATTTGAGAACCATATAAAGTCGACCAAGAATTTCCAG	6062
Db	11164	GCAACCAACACGCTGATCATATGTGAGACACCCCGCAACAAAAATGACCAAGAAATTTCCAG	11222
QY	6063	CGGAGATTTTCCAAAACAATCTTGGAACTGCTGCTTTCACCTGTTTGGGGAGCATATCC	6122
Db	11224	CCGCAATCTCAAAAACTTCAATGAGTTGGCTGTTTTCCTTTTCCGCGGCGCTGCTGCG	11283
QY	6123	TCATTTGTTGAGACTTATATGTGTGGTCTGCAAGCTCTATGCTTATAAACACAGTAAAT	6182

Db 11284 TATTAAATTAAGACTTAATGATTTTCTTCGAGCATGATGCTGACACACGAGAT 11343
Qy 6183 GACTAGCGCGGACACTGA-CATAGCGGTAAACTCGATGACTTCCGAGAGCGTGT 6241
Db 11344 GACCGCTACGCCCATGATGACCGGACGACGAAACTCGATGACTTCCGAGAGCGTGT 11403
Qy 6242 GCATATGCGCAGCGCGCT-----TGACACT 6268
Db 11404 GCATATGCGCTGATGATTAATGATCCCGCTTACCGCGGCAATATGACACACC 11463
Qy 6269 AAAATCGATGATTTCCGAGAGACACAGTGCATTAATGCTGTGAGTGT 6318
Db 11464 AAAATCGATGATTTCCGAGAGAGCGAGTGCATTAATGCTGTGAGTGT 11513

RESULT 8
US-09-367-764-1
Sequence 1, Application US/09367764
Patent No. 6583121
GENERAL INFORMATION:
APPLICANT: Johnston, Robert E.
APPLICANT: Davis, Nancy L.
APPLICANT: Simpson, Dennis A.
TITLE OF INVENTION: System for the In Vivo Delivery and
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
STREET: 1211 East Morehead Street
CITY: Charlotte
STATE: No. 5583121th Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09367,764
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/801,263
FILING DATE: 19-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-147
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11663 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 60..7559
FEATURE:
NAME/KEY: CDS
LOCATION: 7608..11342
US-09-367-764-1

Query Match 36.9%; Score 2396; DB 4; Length 11663;
Best Local Similarity 64.6%; Pred. No. 0;
Matches 3751; Conservative 0; Mismatches 1985; Indels 74; Gaps 9;

Qy 580 ACTGACGGTATGAAACGGGAGCGTATATTTCTCATCGGAAACAGCCAGGTCACTTC 639
|||||

Db 5707 ACTGCTAACCGGGGTAGTGGGTACATATTTTCGACGGAACAGGCCCTGGGCACCTTGC 5766
Qy 640 AAAGAATTCAGTACGTCAATGTAACTAAGAACCATTATGATGGGCCCGCCCATG 699
Db 5767 AAAGAATTCAGTACGTCAATGTAACTAAGAACCATTATGATGGGCCCGCCCATG 5826
Qy 700 AGAAGTATTAAGCCCGCGCGCTCGATCTCGAAGAGAAATGTTACAGAAAGAACTGC 759
Db 5827 AAAGAATTCAGTACGTCAATGTAACTAAGAACCATTATGATGGGCCCGCCCATG 5886
Qy 760 AATTATGCGCTCTGAAGAAATTAAGACAGTATCAATCAAGAAAGTAAATATGA 819
Db 5887 AGATGATCCGACGGAAGCCAAAGACAGTACAGTCTGAAAAGTAAAGAAACCGA 5946
Qy 820 AAGCAATTAAGCGGAGGAGTCAATTTCTGATGGGACATATCTATCAGAAAGGA 879
Db 5947 AAGCAATTAAGCGGAGGAGTCAATTTCTGATGGGACATATCTATCAGAAAGGA 6003
Qy 880 ATCCGTGAGTGTATACAGAGTCAATTAATCTGTAACCAATCTACTGTCAGAGTATTA 939
Db 6004 ATCAAGCAATTCATTAATGATCACTACCGGAACCAATGTTTCCAGAGTATCCAG 6063
Qy 940 ACAGGTTTACATCTGACAGAGGTGCGGTTAAACGTGCACTTAATCCAGAAAGAT 999
Db 6064 CGAATCTACTGACCCAAAGTTGCTGTGAGCTGTTGTAACAATCTATGATGAGAT 6123
Qy 1000 ACCCTACAGTACGAGTATTTATTAACAGTAAAGATAGCGTATCTTACATGCTG 1059
Db 6124 ACCGACGAGTACATCTTATGATACAGTACCGAGTACAGTACGATTAATGATGATG 6183
Qy 1060 ACGGCGCATCGTGTCTGATGATACAGCACTTTTTCGCGGCTAACTGAGAAAGTACC 1119
Db 6184 ACGGACAGTACGCTGTGCTGATATGATGATGATGATGATGATGATGATGATGATG 6243
Qy 1120 CAAGAAGCATAGCTATTTGAGCGAGGATGATGATGATGATGATGATGATGATGATG 1179
Db 6244 CGAAGAAGCATAGCTATTTGAGCGAGGATGATGATGATGATGATGATGATGATGATG 6303
Qy 1180 ATACATTAACAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1239
Db 6304 ACAGGTTGCAAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 6363
Qy 1240 GAGATTAACCTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1299
Db 6364 GTGAATGCCCAACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6423
Qy 1300 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1359
Db 6424 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6483
Qy 1360 TTACGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1419
Db 6484 TTACGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6543
Qy 1420 ATTAATCTAAACCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1479
Db 6544 ATTAATCTAAACCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6603
Qy 1480 ATGTCAAAGTATCTCCGCGACGAAACATACAGAGAGCGGCTTAAGTGCAGGTTATTC 1539
Db 6604 ACGTGAAGTATCACTGCGACGAAACATACAGAGAGCGGCTTAAGTGCAGGTTATTC 6663
Qy 1540 AAGCTGAGATCCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1599
Db 6664 AAGCTGAGATCCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6723
Qy 1600 GACTGAATGCGGCTCTGCGCAATATCACTCTCTGATGATGATGATGATGATGATGATGAT 1659
Db 6724 GCTTAAAGCGCGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6783
Qy 1660 TTGATGCGATATTTGCTGAACATTTCCACAGCGGCGACAGTATTTGAAACGACATTCG 1719
Db 6784 TTGATGCGATATTTGCTGAACATTTCCACAGCGGCGACAGTATTTGAAACGACATTCG 6843

QY	1120	IGTGTGTTTGAATAAAGGGAACACGACGCTATCTGGCACTTTGGCGTGAATGATCTCTAGG	1779
Db	6844	CTTGTTTGCACAAAGGCCAAGACGACGCTAATGGCGTTAACGGCGCTGATGATCTTGGAG	6903
QY	1780	ACTTAGGTGTGACCAACCGGCTCTTAGATTGTATAGAGGCGGCGTTCGGCAATATCAAT	1839
Db	6904	ACCTGGGTGTGACCAACCACTACTCGACTTGATCGATGGCGCTTTGGAGAAATATCAT	6963
QY	1840	CTGTGCACCTTACTACAGGAACGAGGTTAAATTGGTCCATGATGAAATCCGGTATGT	1899
Db	6964	CCACCCATCTGCCCACGGGTACCCGGTTCAAATTGCGGGCGATATGAAATCCGGAATGT	7023
QY	1900	TCTTAACGCTGTTTGTCAACACACTAGTCAATATCATATGATATTGCTAGCAGATCAAGTG	1959
Db	7024	TCTCTACGCTCTTTGTCAACACAGTTCGAAATGTGTATTATGCGACAGAGATATTGAGG	7083
QY	1960	AACGGTTAACCACTGACGCGTGGCGGCGCTCTATGGCGACGATPACATAGTCATGGTG	2019
Db	7084	AGCGGCTTTAAACGTCMAATGTGACGATTTATTCGGGAGACGACAACTTTATACGGAG	7143
QY	2020	TCTGTCTCGACACCTTGATGGCGGAGATGGCGCACTTGGCTGAATCGAATGGAATGTA	2079
Db	7144	TAGTATCTGACAAAGAAATGGCTGAGAGGTGTCCACTGGCTCAACATGGAGGTTAAG	7203
QY	2080	TTATTGATGACGTTATTGTATCAAAGCACCTTACTTGTGTGGGGGATTTATCTGTGG	2139
Db	7204	TCATTGAGCGAGTATCGGCGAGAGACCACTTACTTCTGCGGTGGATTCATCTTGGCAAG	7263
QY	2140	ACCAAGTATACAGGACGACAGCCTGCGAGAGTCGAGACCTCTTAAAGGCTTTTAAAGCTG	2199
Db	7264	ATTGGTTACTCTCAACAGCGTGTGCGGTGGGAGCCCTTBAAGAGGCTGTTTAAAGTGG	7323
QY	2200	GAAGAACCATTCGCCAGTCGATGATACCAAGACTGGACCGCGCGCGGACCTGCATGATG	2259
Db	7324	GTAACCGGCTCCACGCGCAGCAGTACGAAGACAAAGACAGAAAGACGGCTCTGGCTAGATG	7383
QY	2260	AAGCAATTCGATGGAACGAATTGGAAATTACGACGAGTGTGTAGAGGCGGTAGATCCA	2319
Db	7384	AAACAAAGCGCGTGTTTAGAGTATGATTAACAGACACCTTACAGTGGCCGTGGCAATCTC	7443
QY	2320	GATACGAGATCATACTGGCAGGCGCTGCATCATCAAGTCTCTGACCGACGTTACCGAAAGCG	2379
Db	7444	GGTATGAGGTGACACACATCACACTGTCTCTGTGCAATTAGAACTTTTGGCCAGAGCA	7503
QY	2380	TTAAGAACTTCAAGAGCATTAAGGAGGAGCCCAATCACCTCTGACGGCTGACCTTAATAGG	2439
Db	7504	AAAGAGCATTTCCAGCCCATCAGAGGGGAAATTAAGCATCTCTACGGTGTCTTAAATAGT	7563
QY	2440	TGAGGTAGTAGA-----CAGCAGACTTACCCACCGGAGAAATGTTTCCA	2482
Db	7564	CAGCATATTAACCTTTCACTGTGATCTAATACACACACACACACCACTGATAGAGATTTCT	7623
QY	2483	TACCTCAGCTGAACCTTTCACACAGTTTACCTTCAAAATTCGATGGCTTACCGAGATCCA	2542
Db	7624	TTAATCATCTCGGCGCGCCCTTTCCAGCCCCCATCTGATGAGGCGCGGAGAA	7683
QY	2543	AACCTCTTAGCGCGCGCTGAGAGCGCTTTGGCGCCCGCTGTGCTCAATGGAAGAT	2602
Db	7684	GAGGCGAGGCGGCCCCGATGCGCTCGCCGCAATGGGCTGGCTTCCAAATCCAGCAACTGA	7743
QY	2603	CTTAGAGAGTGGATGTGATCACTTGACTTTCAAAACAC--GATTCACCTTAATCCGCGCAG	2660
Db	7744	CCACAGCGGTGAGTCCCTTAGTCAATTGACAGGCACTAAGACTTAACCCCAAGCCAC	7803
QY	2661	GTCACCGCCAAAGAAAGAAAGAGTGTCTCTAAGCCAAACCTTACTGAGCTTAAAGAA	2720
Db	7804	GCCGCGCCCGCGCCAGAAAGACGAGCGCCAAAGCAACACACGAAAGCCGAAAGAAACAA	7863
QY	2721	AGAGGACACCAAGCAAGAGCAAGAAC---GCAAGCTTAAACAGAGGAAAGACGACACTA	2777
Db	7864	AAACACAGAGAAAGAAAGAAAGCAACTGCAAAACCCAAACCCGGAAGAGACAGCGCTA	7923

QY	2778	TTGTGATGAAGTTGAGATCGGACAAAGCATTTCCGATC---ATGCTGAAGCGGCCAACTGA	2834
Db	7924	TGGCACTTAAATTGTGAGGCGCAGACTGTTCGACGTCTAAAAATGAGAACGGAGATGTCA	7983
QY	2835	ATGATATATGCTCGTGCGTTGTTCGAGAGAAAGCGCTGATGAACAACCTCCAGTTGAAGAAAA	2894
Db	7994	TCGGGCAAGCATGCGCCATGGAAGGAAGATATGAAAACCACTCCACTGTAAGGAAGACTA	8043
QY	2895	TTGATTAATGAGCAATTAGCGGCCGTGAATTGAAGAGCTAGCATATGACCTTGGAGT	2954
Db	8044	TTGACCAACCCCTGTGCTATCAAAAGCTCAAAATTCAACCAAGTGGTACGTATACGACATGAGT	8103
QY	2955	ACGGCGACGTTCCCGCAAGACATGAATATCAGACAGCTGACAGTATACACGAGCAAAACAC	3014
Db	8104	TCGCACAGTTGCGCGGTCAACATGAAAGTAGAGGGTTTCACTTACCAAGATGAACACCTTG	8153
QY	3015	CGGGCTTTCAACATGCGACCAACCGGCGCAGTCCAGTATGAGAAATGGAGATTTTACCGTAC	3074
Db	8154	AAGGTTTCAACATGCGACCAACCGGCGCAGTATGAGAGCGAATTTACATCC	8223
QY	3075	CGAAGAGTGGGCGGGAAGGCGACAGCGGAAGCCGATCTTGAGAACAGAGGCGAG	3134
Db	8224	CCCGCGAGTAGAGGCGAGAGAGACAGTGGTGTGTCGATATGATATACACAGCGCGGG	8283
QY	3135	TTTGCGCTATTGTTCTTAGAGAGTCAATATGAGGCAACCGCTTACGCGCTTCACTGCTCA	3194
Db	8284	TTTGTCGCATATGCTCTCGAGAGGGCTGATAGAGGAACAAGAACCGCTTTCGGTGTCTCA	8343
QY	3195	CTTGGAACCAAGAAAGGGGTGACCAATTAGAGGATACCCCGAAGGTTCTGAACCGGTGT---	3251
Db	8344	CCTGGAAATGCAAAAGGAGAGACAAATCAAGACAAACCCCGAAGGGAACAGAAAGTGTCTG	8403
QY	3252	-----CACTAGTTACAGCGCTATGCGTGTCTTTCGAAATGTCACTGTCCATGCGACAAAC	3305
Db	8404	CTGCACCACTGCTCAGCGCCATGTGCTTCTTGGAAACGTGACCTTCCCATGCAATGCACC	8453
QY	3306	CACCCGTGTGTAATTTACTAGACGCCGACAAAGCAACATCTCGACGTCTGAAAGAAACCTCG	3355
Db	8454	CGCCCACTGTCTAACCCGCGAACATCTCAAGGCTCTCGACATCTTGAAGAGAACTGA	8523
QY	3356	ACAATCCAAATTACGACACGCTGCTGGAGAACGTCTTGAAT---GTCCATACAGCGCGGC	3422
Db	8524	ACCAAGAGGCTTACGACACCTTGTCTCAACGCAATTTGCGGTGCGGATCTGTCCGCGAGAA	8583
QY	3423	CCAAACGAGCATTAACCGATGACTTCACTGACAGTATCCCTACCTGSGGGTTCTGCCGT	3482
Db	8584	GTAAGAAGAGGTCACTGACGACCTTTTACCTTTGACCAAGCCGTAATTGCGGACATGTCTGT	8643
QY	3483	ATTGCAGACACTCAACGCCGCTGTTTCAGCCCAATATTAATTTGAGAACGTGTGGAGCAAT	3542
Db	8644	ACTGCACCAATACATGAAACCGGCTTTTAGCCCGATTTAAGATCGACAGAGTCTGGGATAGG	8703
QY	3543	CTGATGATGATCGATTGAAATCCAGGCTCTGGACAAATTCGCGCTAACATCAGGACAGCA	3602
Db	8704	CGAGACGACAAACCAATACGATACAGACTTTCGCCCACTTTGGATATGACCAAGCGGAG	8753
QY	3603	CTGCGGATGTCAACCAAAATTCGGTTACATGTCTTTGACACAGCAACCATGACATCAAGAGG	3662
Db	8754	CAGCAGAGTCAAAATTAAGTACCGGCTAATATGTGCTCGAGCAGAGATCATATCTCAAGAGG	8823
QY	3663	ACAGTATGAGAAATATAGCTATACGACATTTGGAACCTCGCGCTGTCTTGGCAAAAG	3722
Db	8824	GCACCATGATGACATCAAGATTCAGACCTTCAGAGCCGTGTAGAAAGGCTTAGCTTACAAAG	8883
QY	3723	GGTACTTCCTGTAGCTCAATGTCTCTCAGGTGACAGTGTAAACCGTACAGTATCAAGGCG	3782
Db	8884	GATATCTTCTCTCGGAAGTGTCTCCAGGGGACAGGTAAACGGTTAGCATACGAGGTA	8943
QY	3783	GAGCATCTGAATTTATGACACCGTGGAGAAAAAGATCAGAGAGAGATTTGTTCGTGAG	3842
Db	8944	GCACTCTGACAACTGATGACAAATGTGCGCCGCAAGATTAACCAAAATTCCTGTGGAGCGGG	9003
QY	3843	AGGAGTACTTGTCCACCGCTGCATGGAAGCTGTAAAGTCCACAGCTTTACATCACT	3902

Db 9004 AAAAATATGACCTTACCTCCGTTTCAAGGTAAGAAATTCCTTGCACAGGTGACACCCGTC 9063
Qy 3903 TGAAGAGACGTCCTCCGGGTATCATATCCATGACAGGCGCCACACGGGTATAGT 3962
Db 9064 TGAAGAAACAACCCCGGTATCATATGACAGGCGCCGACCGCATGCTTATACAT 9123
Qy 3963 CCTATCTGGAGAGACGTCAGGCGAAGTATCATTTAAACCACTTCTGGCAAGACGTCA 4022
Db 9124 CCTATCTGGAGAGATCATAGGGAAGTTTACCGAAGCCATTCGGGAAGAACATTTA 9183
Qy 4023 CCTATGAAATGTAAGTGTGCGCATACAGACAGTATCTGAGACCGCAAGCAAGATGA 4082
Db 9184 CGTACGAGTGAAGTGGCGCGATTCAGAACCGGAACCGTTACGACCCGTACCGAAATCA 9243
Qy 4083 ACGGTGACCTTAACCAAAACGATGCTTGTCTTAAGAGGCAACCAAGCAAAATGGGCT 4142
Db 9244 CGGGCTGACCGCCATCAAGCAGTGTGCTCTTATTAAGCGACCAAGAAAGTGGGCT 9303
Qy 4143 TCAACTCGCCGGATCTTATTAAGGACACAGACCACTAGTGAAGTAAATTTGACATTC 4202
Db 9304 TCAACTCGCCGGATCTGATCAGACAGCGCCAGCAACGGCCCAAGGAATTTGCTTTC 9363
Qy 4203 CATTCGGCTTGACACGACAGTCTGCGCGTTCCGTTAGCTTCAACGCGCTACAGTCA 4262
Db 9364 CTTCAGGTGATCCGAGTACCTGATGTCTCTGTTGGCCACGCGCGCAAGTATAC 9423
Qy 4263 AGTGTTCAGAGGATCACTCTTCACTGATCTGCAATGAGCAACATTTGCTGACAGCA 4322
Db 9424 ACGGCTTTAAACATCACTACGCTTCAATTAAGACAGACCATCTGACATTTGCTACACCA 9483
Qy 4323 GAAATTTGGGGGTGAGAGACGCAACGAGAAATGATTAAGAGGTCTACATCCAGCA 4382
Db 9484 GAGAGCTAGGGGCAACCCGGAACCACTGAATGATCATTCGAAACACGGTTAGAA 9543
Qy 4383 ATTTTCTGTGGGGGAGAGGGCTGAGTACGATAGGGTAAACATGAACCATGCAAG 4442
Db 9544 ACTTACCGTGCACGAGATGCGCTGGAATACATATGGGGCAATCAGAACCATGAAGG 9603
Qy 4443 TCTGGGCCCCAGAGTCCGACCAAGGCGACCCCATGATGCGCGATGATCATCATTC 4502
Db 9604 TCTATGCCAAGTCTGACAGAGAACCTTCAAGATGCGCAACGAAATGATACAGC 9663
Qy 4503 ACTATTTATCATCGGATCCAGTCTACCTGATGTGTGTGTGTGTGTGTGTGTGTGT 4562
Db 9664 ATTACTATCATGTCCATCTGTGTACCATCTTACCGCTCGCATCAGTGTGTGTGTGT 9723
Qy 4563 TCTGTGAGGACCTGATCATCAGAGCTTGCATGCGCAAGCAAGAAAGACTGCTGA 4622
Db 9724 TGATGATTTGGCGTAACTGTGTGACATTAATGTGCTGTAAAGCGCGGTGAGTCTGA 9783
Qy 4623 CGCCATAGCGGCTTGTGACAGCAAGCGGATCCCAAGCATTTAGCGGTTTGTGTGTGA 4682
Db 9784 CGCCATAGCGGCTTGTGACAGCAAGCGGATCCCAAGCATTTGTGTGTGTGTGTGT 9843
Qy 4683 TTTCGCAACCAAGCGTGAACATTTGAGAACTTTGAAACATCTGTGTGTGTGTGTGT 4742
Db 9844 TTAGGTGGCTTAATGTGAACATTTGAGCAAGCATGATTTAATTAATGTGTGTGTGT 9903
Qy 4743 AACCGTTTCTGTGGACAGTGTGATCTCTGTGACGCGTTGTATTTCTGTGTGTGT 4802
Db 9904 AGCGTTTCTGTGGTGTGACAGTGTGTATCTCTGTGCGGTGTGTGTGTGTGTGTGT 9963
Qy 4803 GCTTTTATGCTGATGCTTTTATTTATTTGTGTGAGGCGGTGTGTGTGTGTGTGTGT 4862
Db 9964 GTTGTCTATGTGCTGCTTTTATTTATTTGTGTGTGCGGCGCTTACTGTGTGTGTGT 10023
Qy 4863 CTTTGAACATGAGCACTGTGCAAAATGTTCCGGGATTCGGTATTAAGGCTGTGTGT 4922
Db 10024 CTTTGAACATGAGCACTGTGTCAAAATGTGCCACAGTATCCGATTAAGGCACTTGTGT 10083
Qy 4923 AACGCGAGGTTACCGGCACTTAACTGTGAGATCAGGTCGTCTCATCGAATTAAC 4982

Db 10084 AAAGGCGAGGGTACGCCCCGCTCAATTTGAGATTAATGTCTGATGTCTCGAGGTTTTC 10143
Qy 4983 CTTCAACTAACAAGAGTACGACCTGCAATTCACACAGTCAATTCCTTACACCAAG 5042
Db 10144 CTTTCAACAACAAGAGTACATTAATGCAAAATTTACCACTGTGTGTGTGTGTGTGTGT 10203
Qy 5043 TTAATGTGTGGGTCTCTGAGTGCAGAGGATCTTCAAGGCGGATTAACATGCGCG 5102
Db 10204 TCAGTGTGTGGGTCTTGTGAATGTGACCGCGCGCTCAGGACATTAATCTGTCAAG 10263
Qy 5103 TTTTGGCGGTGTGATCCCTTTATGTGGGAGCGCAATGCTTGTGTGACATGAGA 5162
Db 10264 TCTTTGAGGGGTGTATCCCTTCAATGTGGGAGGAGCAATGTTTGTGCGACATGTAGA 10323
Qy 5163 AACACAACTAGTAGGCGGTACGTGAGTGTGTGCGACATGCACTATTAATCAAGAG 5222
Db 10324 AACGCAATAGTAGGCGGTACGTGATTTGTGATGATTTGCGGACTGACACGCGC 10383
Qy 5223 TCGCACTAAAGTTTACACAGCTGCTGAAAGTGTGCGCTGCTATTAATATGCAACA 5282
Db 10384 AGCGATTAAGTGTGATCTGCGGAGTGAAGAGTACGCTGCTATGATGTGCGGAACA 10443
Qy 5283 CCACCGCGACCTTGATACGTTTGTCAATGAGCGTACGCAAGTTCCTACGGGACTGA 5342
Db 10444 CTACCAAGTTTCTAGATGTGATGAGGAAGAGTCAACAGGAACGCTTAAAGACTGA 10503
Qy 5343 AGCTATAGCAGGCGGATATCAGCGGCTTTTACACCTTTGACATTAAGTGTGATCA 5402
Db 10504 AAGTATAGCTGAGCAATTTTACATGATTTTACACATTTGATCAAGGTCGTATCA 10563
Qy 5403 GAAAGGGGCTTTTACACTACGACTTCCGTGATAGAGCTATGAACCAAGAGGCT 5462
Db 10564 ATCGGCGCTGTGTGACATCTGATCTTTCCGGAATAGAGGACATGAACCAAGAGCT 10623
Qy 5463 TCGGATATTTCAAGCATCTCGCTTGTATGCTACAGACATAGTACCGGCACTGATAC 5522
Db 10624 TTGAGACATTTCAAGCTACTCTTGACTACCAAGACCTCATGCGCAGCAGACATTA 10683
Qy 5523 GGTCTGTAAAGCTTGTGTCAAGAACATCCAGCTTCCCTTAACCAACAGATTAAGGT 5582
Db 10684 GGTACTCAAGCTTGTGTCAAGAACATCCAGCTTCCCTTAACCAACAGATTAAGGT 10743
Qy 5583 ATGAATGTGAAAGAACACTCAGAGCAAGCCCTGCAAGAAACAGCACTTTGTGATGA 5642
Db 10744 TCGAGATGTGAAAGAACACTCAGGCGGCACTGCAAGAAACCGCTTTGTGTGTGA 10803
Qy 5643 AAATTAAGTGAAGCTTGTGTGAGGCTTAACTGTGTACGAGCAATCTCTATCTGA 5702
Db 10804 AGATTGACAGTCAATCCGCTTGAAGGCTGACTGTCTATCGGGAACATTCATTTCTA 10863
Qy 5703 TTGACATCTGTGATGAGCTTTTGTGAGATCATCAGAAATCAACAAATTTTGAAGTTA 5762
Db 10864 TTGACATCTGTGAGCTTGTGTGAGAGATCATGAGACACTGTGTCTCAACAGTCA 10923
Qy 5763 GCTGACAGTGAAGACTCATTTATTTGAGACTTGTGTGAGCTTGTGTGATCATTAAGT 5822
Db 10924 AATGTGATGTAGTAGTGTGACTTATTTAGCGGACTTGTGAGGATGTGACTGTGAGT 10983
Qy 5823 ACAAGTGAAGAGGAGGACATTTGTCCAGTTCACTCCCATCCACGACAGCTGTTTGA 5882
Db 10984 ATGTATCCGACCGGAAGGACAAATGCTGTATCAATTTGACATTTGACAGAACCTTCC 11043
Qy 5883 AGAAGGACCAACATGTGACTGCGTGAAGCAGATTAACCTATTTTGTGACATGGA 5942
Db 11044 AAGGTGACAGTTCAATCTCTGAGAAAGAGCGGTGACAGTACACTTGTGACACCGGA 11103
Qy 5943 GCCCAAGCAAAATTTATTTATTTGTGATGCGGCAAGAGTCAACCTGCAATGCTGAT 6002
Db 11104 GCCCAAGGCAATTTATTTATTTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11163
Qy 6003 GTAAACCAACCGGCGACCACTTAATTTGAGAAACCACTAAAGTGCACCAAGATTTCCAG 6062
Db 11164 GCAACCAACGAGTGTATGTGTGTGACACCCCGCAAAATTAAGCAAAATTTCCAG 11223

Db 6883 CATCATTCGACAAAAGCCAAAGCAGACGCTATGGCGCTTAACCGGCTGATGATCTTGAGG 6942
Qy 1780 ACTTGGTGTGACCAACCGCTTATAGATTGATAGAGCGCGCTTCCGCAATATCACT 1839
Db 6943 ACTGGGTGTGATCAACACATCTGATCTGATGATGAGCGCTTGGAGAAATATCAT 7002
Qy 1840 CTGTCCACCTACCTACAGGAAGAGGTTAAATTGGTCCATGATGAAATCCGGATGT 1899
Db 7003 CCACCATCTACCTACGGGTACTCGTTTAAATTGGGGCGATGATGAAATCCGAAATGT 7062
Qy 1900 TCTTAACGCTGTTGTCAACACACTAGTCAATATCATGATTCTAGCAGATTAACGTG 1959
Db 7063 TCCCAACCTTTTGTCAACACAGTTTGAATGTGGTATCCGACAGAGTACTAGAAAG 7122
Qy 1960 AACGTTAACACGCTCAGGTCGGCGGCTCTATCGGCGACATTAACATAGTCATGTG 2019
Db 7123 AGCGCTTAAACGCTCAGATGTGACGCTTCAATTGGCGACGACCAATCATATCATGAG 7182
Qy 2020 TCGCTCCGACACCTTGTATGGCGGAGATGCGCCACTTGGCTGAACATGAGTAAATA 2079
Db 7183 TAGTATCTGACAAAGAAATGGCTGAGAGGTGGCGCACTCGCTCAACATGAGGTTAAGA 7242
Qy 2080 TTATTGATGACATTATGTTGATCAAGCAACCTACTCTGTGGGGATTTATCTGTGTG 2139
Db 7243 TCATCGACGACATCGGTGAGAGACACCTTACTTCTGGCGGATTTATCTTGCAAG 7302
Qy 2140 ACCAATTAACAGGACAGCTTCGACAGTGTGACAGCCCTTAAAGGCTTTTAAAGCTTG 2199
Db 7303 ATTGGTTATCTTCAACAGGTCGCGGTGGCGGACCCCTGAAGGCGTGTAAAGTTAGTTGG 7362
Qy 2200 GAAACCAATTGCGATGATTAACCAAGACTGCGACCGCGCGGCACTGATGATG 2259
Db 7363 GTAAACCGCTCCAGCCGACGACGAGCAAGACGAAAGACGAGCGCTCTGCTAGATG 7422
Qy 2260 AAGCAATCGATGAAACAGAAATTGGAATTACGAGAGATTAGTGAAGCGCTAGAAATCA 2319
Db 7423 AAACAAAGCGGTGTTAGATGATGATTAACAGGCACTTACAGTGGCGGTGACGACCC 7482
Qy 2320 GATAGAGATCTATCTGGCAGGCTGATATCACTCTCTGTCAACGTTAGCCGAAAGCG 2379
Db 7483 GGTATGAGGTAGACAATATTACACCTGTCTCTGCTGATTTGAAACTTTTGGCCAGAGA 7542
Qy 2380 TTAAGACCTTCAAGGACATTAAGAGGAGCCCAATCAACCTTACGGCTGACCTAATAGG 2439
Db 7543 AAGAGCAATTCBAAGCATTCAAGGAGGAAATTAAGCATCTTCAAGGTCTCTAATAGT 7602
Qy 2440 TGACGTAGTAGACAGCACTTAC-----CCACCGGCAAGATGTTCCATACC 2486
Db 7603 CAGCATAGTACATTTCACTGACTAATCTACACACCAACCATGATTAAGATTTCT 7662
Qy 2487 CTCAGCTGAATTTTCAACCACTTTTACCTCAAAATCCGATGCTTACCGATCCAAACC 2546
Db 7663 TTAAATGCTGCGCGCGCGCTTCCGCGCCCACTGCTCATGTGAGGCGCGCGAGAA 7722
Qy 2547 CTCCTAGCGCGCTGAGGCGCTTTCGCGCCCGCTGCTGCTCAATTCGAAGATCTTA 2606
Db 7723 GAGGACAGCGCGCCCGATGCTTCCCGCAACGAGGTGCTTCTCAATTCACGAATCTGA 7782
Qy 2607 GAGAGTGAATGATCACTTCACTTCAACCAAGTCACTAATCCGCGCGAGGTGCA- 2665
Db 7783 CCACAGCGGTGAGTCCCTAGTCAATTGACAGGCAACTGACCTCAACCCCACTCCAC 7842
Qy 2666 -----CCGCGCAAGAAAGAAAGTGTCTTAAAGCAAACTTACTACCTTAAAGAA 2720
Db 7843 GCCCGCAACCGCGCGAGAGAGAGCGCGCCCAAGAACCAACGGAAGCGGAAGAACCA 7902
Qy 2721 AGAAGCAGCAAGCAAGAGAGAGAAAC---GCAAGCTTAAACCAAGAGAAACGACAGTA 2777
Db 7903 AAACGAGAGAGAGAGAGAGAGAACTGCAAAACCCCAACCCGAGAAAGACAGCGGA 7962
Qy 2778 TGTGTATGAGTTGAGTCTGACAAAGACATTTCCGATCATG---CTGAACGCGCAAGTGA 2834

Db 7963 TGGCATTAAAGTTGAGGCCGACAGATTTGTCAAGCTCAAGAACGAGACGAGATGTCA 8022
Qy 2835 ATGATATGCTCGCTTGTGAGAGAGGCTGATGAACCACTCCAGCTTGAAGAGAAA 2894
Db 8023 TCGGCGACGACTGCGCATGAGAGAGAAAGTAAAGTAACTCTGACGTGAAAGAACCA 8082
Qy 2895 TTGATATGACCAATTAAGCCGCGGAAATGGAAGAGGCTAGCATGTACGACTGGAGT 2954
Db 8083 TCGACACCCCTGTGTATCAAAAGCTCAAAATTTACAAATGCTGACATATACGATGTAGT 8142
Qy 2955 ACGGCAGCTTCCCGAAGACATGAATACACACGCTGACATGACACGAGCAAAACAC 3014
Db 8143 TCGCAGATGCTGACATGACATGAGAGAGTGAAGGCACTTACCTACCACTGACGAGACCCG 8202
Qy 3015 CGGCTTTTAAACCTGAGACCAACCGCGGACATCTCAATATGAGAAATGGAGATTTTACCCTAC 3074
Db 8203 AAGATTTATTAATCGGACCAACGAGGCGGTGACATATGAGAGTGAATTTTACATTC 8262
Qy 3075 CGAGAGAGTGGCGGAGAAAGCGACACGAGAAAGCCGATCTGGAACAAAGAGGACAG 3134
Db 8263 CTCGGGAGTGAAGAGCAGAGAGACAGCGGTCTGCTGATATGATTAATCTCCGTCGG 8322
Qy 3135 TTGTGCTATTGTTCTAGAGGTGCAAAATGAGGACGCGTACGCGCTTTCAATGTCTCA 3194
Db 8323 TTGTGCGATAGTCTCGGTGAGGCTGATGAAGAACGAACTGCTTTCGTCGTCTCA 8382
Qy 3195 CTGGAACCAAGAGGAGTACATTTAGGATACCCCGGAAGTTCTGAACCGTGTGT--- 3251
Db 8383 CCGGAATGATTAAGAGGAGAACAAATTAAGACGACCCCGGAAGGACAAAGAGTGTGCG 8442
Qy 3252 -----CACTAGTTACAGCGCTATGCGCTTTTGAATGTCAACGTTCCATGCGCAAA 3305
Db 8443 CAGACACCTGTGACGCGCAATGTGTCTGCGAAATGTAGCTTCCATGCGACCGCC 8502
Qy 3306 CACCGTGTGTTACTGACGACCGCCAGAACGAACTGACGCTGTGCAAGAGAGCGTGC 3365
Db 8503 CGGCCACATGCTATCCGCGAACCTTCAGAGCCCTGCAATCTTGAAGAGACGTGA 8562
Qy 3366 ACAATCCAAATTAAGACACGCTGTGAGAAACGTTTGAATGTCTATCAC---GCCGC 3422
Db 8563 ACCATGAGCGCTTACGATCCCTGTCAATGTGCGATGCGATGCTGTGCAAA 8622
Qy 3423 CCAAGCAAGATTAACGATGATCACTGACAGTCCCTACCTGGGATTTCTCCCGT 3482
Db 8623 GCAAAAGAGGCTCACTACGACTTTTACCTGACAGCCCTTACTTGGGACATGCTGT 8682
Qy 3483 ATTGCAACACTCAACGCGTGTTCACGCCCAATTAATTAATGAGAACGTGTGAGCAAT 3542
Db 8683 ACTGCACCATATGAAACGATGCTTACGCTTCAATGATGAGAGAGGTCTGGAGCGAAG 8742
Qy 3543 CTGATGATGATCGATTGAATCCAGGTCTGGGCACAATTTGGCTTCAATCAGGAGCA 3602
Db 8743 CGGAGATTAACCATTAAGCATGACACTTCCGCGCCAGTTTGAATCAACCAAGCGGAG 8802
Qy 3603 CTGCGGATGACCAAAATTCGTTTACATGTCTTTTGAACAGACCATGACATCAAGAG 3662
Db 8803 CAGCAAGCGCAAAAGTACCGCTACATGTGCTTGAACAGATACACCTTTAAGAGAG 8862
Qy 3663 ACAGATGAGAAATATGACTATGACACATGTGACACCTGCGCTGTGCGCCCAAG 3722
Db 8863 GCACCATGATGATCAAAATTAAGACCTCAGGACGCTGTAAGAGGCTTATGCTCAAG 8922
Qy 3723 GGTACTTCTGTAGCTCAATGTCTCTCAGGTGACAGTGTAAACGTCATACAGAGC 3782
Db 8923 GATACTTCTCTCGCAAAATGCTCTCAGGAGGACAGGTAAACGTTAGCATGTAGATA 8982
Qy 3783 GAGCATGAGAAATTAATGACACCGTGAAGAAAGATCAGAGAGAGAGTGTGCTAGAG 3842
Db 8983 GCAACTGCAACCTGATGTACATGCTGCGCCCAAGATTAACCAAAATTCGTGGACGG 9042
Qy 3843 AGAGTACTTGTCTTCCACCGCTCATGAGAAAGCTGTAAAGTGCACCTTTACATCACT 3902
Db 9043 AAAATATGATCTACCTCGCTTACAGGTAAATAATTCCTTGACAGTGTACGACGCTC 9102

QY 3903 TGAAGAGAGCTCTGCGGGTACATTAACATGACAGGCCAGGCCACACCGCTATAAGT 3962
DB 9103 TGAAGAAACCACTGAGGCTACATCACTATGACAGGCCAGGCCACCGCTATATCAT 9162
QY 3963 CCTATCTGAGAGAGCGCTCAGGCGAAGTGTACATTTAAACCACTTCTGGCAAGACGCA 4022
DB 9163 CCTATCTGAGAGAGTTCATCAGGGAAGTTTACGAAAGCCGCCACTTGGGAAGAACATTA 9222
QY 4023 CTATGAAATGTAAGTGTGGCACTACAGACAGAGTATCTGAGCAACGGAACGAATGA 4082
DB 9223 CGTATGAGTGCAGATGGCGGACATCAAGACCGGAACGTTTCCAGCCGCCAATAATCA 9282
QY 4083 ACGGCTGCACTAAAGCAAAACAGTGCATTGCTCAAGAGGAGCAACGAAGGGGCT 4142
DB 9283 CTGATTGACCGCCATCAAGACAGTGCCTGCTATTAAGAGGACCAACGAAGTGGCT 9342
QY 4143 TCAACTGCGCGATCTTATTAAGGACACAGACCACTGATGCAAGTAAATTTGACATTC 4202
DB 9343 TCAACTCACCGGACTTATATGACATGACACACACCGGCCCAAGGAATTTGACTTGC 9402
QY 4203 CATTCGCTTGAACACGACAGTCTGCCCGGTTCCGTTAGCTACACGCTTACAGTCA 4262
DB 9403 CTTTCAAGTTGATCCCGAGTACTGACATGCTCCCTGTTGCCACGCGCGAATGTAATAC 9462
QY 4263 AGTGGTTCAAGGATCACCCCTCAACCTGACTGCAATGCCACCAATTTGTCACACGA 4322
DB 9463 ATGGCTTTAAACATACAGCTCTCAATTAAGATACAGACCACTTGAACATTTGTCACCA 9522
QY 4323 GAAATTTGGGCTCTGAGACAGCCACACAGAGATGATTAACAGGCTTACATCCAGGA 4382
DB 9523 GGAATCTAGGGGCAAAACCCGGAACCAACCACTGATGATGCTGGGAAGACGGTCAAG 9582
QY 4383 ATTTTCTGTGGGCGAGAGAGGCTGAGTACGATGGGGTAACTATGACCAAGTCAAG 4442
DB 9583 ACTTCACCGTTCAGCCGAGATGGCTGGAATACATATGGGGAATATCATGACCACTGAG 9642
QY 4443 TCTGGGCGGAGGATTCGACACAGGACCAACCAATGATGGCGCATGAGATCATCATCC 4502
DB 9643 TCTATGCCCAAGATTCAGACACAGAGACCTCTACAGAGGACCAACGAATATGATCAC 9702
QY 4503 ACTATTAATCATCGGATCTCACTGATCTATGCTGCTGTGCTGTGCTCTTCTGCTA 4562
DB 9703 ATTAATCAACATCGGATCTCTGTATACATCTTACCGCTGTGACATCACTCCGTGGGA 9762
QY 4563 TCTGTAGGACATGCACTCATCAGCAGCTTGCACCAAGCAAGAAAGACTGCTGA 4622
DB 9763 TGATGATTTGGGTAACGCTTGAGTATGCTGTAAAGCCGCGCTGAGTGCCTGA 9822
QY 4623 CGGCATACGCGCTTGCACCCGACGCAACGGAACCCACAGATTAAGCGTTTGTGCTGA 4682
DB 9823 CGGCATACGCGCTTGCACCCGACGCAACGCTTGCCTGCTGCTGCTGCTGCTGCTGCG 9882
QY 4683 TTCCGCAACCAACGCTGAAACATTTGGGAACCTTGAACCATCTGTGTTTAAACAC 4742
DB 9883 TTAGTGGGCAACGTGGAACGTTTACCGGACCATGATGATCTGTGTGCAACGTC 9942
QY 4743 AACGTTTCTGTGGACACAGTTGCACTCTCTGGGACGCGTGTATTTCTGTCGCT 4802
DB 9943 AGCGTTTCTGTGGGTCAGTTGCACTCTTGGCGCTTTCATCGTTTAAATGCGCT 10002
QY 4803 GCTTTTCATGCTGATCGCTTTTATTTGTTTGAAGGCTGTGCTGGGGAAGGTAAAG 4862
DB 10003 GCTGCTCTGTGCTGTCTTTTATTTAGTGTGCTGGGCTTACCTTGCGGAAGTAAAG 10062
QY 4863 CCTTGAACATGACACCACTGTGCAAAATGTTTCGGGAGTCCGCTATTAAGGCTGTGCTG 4922
DB 10063 CTTAGCAATGACACCACTGTTCCTTCAATATGCAACAGATACGTTATAGGCACTTGTG 10122
QY 4923 AACGCGAGTTACGCGCACTTAACCTGAGATCACGCTGTCTCATTCGGAATTAACAC 4982
DB 10123 AAAAGGAGGGATATGCCCGCTCAATTTGAGATCACTGTATCTCTCGAGGTTTTCG 10182

QY 4983 CTTCACTAACAGAGATGACGATCGTCAAAATTTCCACAGCTCATCTTCCACACAG 5042
DB 10183 CTTTCAACCAACAGAGTATATTAATCTGCAAAATTTACCACTGTGTGCTCCCTCCCAAAA 10242
QY 5043 TTAATGCTGCGGATCCCTGAGTGCAGGACATCTTCAAGGCGGATTAACATGCGCG 5102
DB 10243 TCAATGCTGCGGCTCTTGTGAATGTGACGCGCGCTCATGCAATATATCTGCAAG 10302
QY 5103 TTTTGGCGGTGTATCCCTTTATGATGGGGAAGCGCAATGCTTCTGTGACAGTGA 5162
DB 10303 TCTTGGAGGGGTCTACCCCTTTATGTGGGGAAGGCAATGTTTGGCAAGTGA 10362
QY 5163 ACACACACTAGTGAAGGCGGTAGCTGAGTTCCTCCAGACTCAGATATAGATCACAG 5222
DB 10363 ACAGCAGATGATGAGGCGGTATGCTGAACTGTACAGAAATGCGCTTGCACACCGC 10422
QY 5223 TCGCACTAAAAGTTTACACAGCTGTCTGTGAAGTGGCGCTGCTGTATATGATACGGCA 5282
DB 10423 AGCGATTAAGTGCACACTGCGCGATGAAGTAAAGTAACTGCTATGTATGAGGGAACA 10482
QY 5283 CCACCGGCACTGTGATACGTTTGTCAATGGCGTCAACGCGAGTTCTCACGGGACCTGA 5342
DB 10483 CTACCACTTTCTAGATGTGTACGTGAACGAGTCAACAGGAACGCTTAAAGACTTGA 10542
QY 5343 AGGTATAGCAGGGCGGATATACAGCCGCTTTTCACTTGAACATTAAGTGTGATCA 5402
DB 10543 AAGTATAGTGCACCAATTTTACGATCTGTTACGCTATTCATATAGTGTGATTC 10602
QY 5403 GAAAGGGCTTTGTTTACACTACAGCTTCCCTGATATGAGAGTATGAAACAGAGCGT 5462
DB 10603 ATGCGGCTGTGTATCACTATGACTTCCGGAATATGAGGATGAACAGAGCGT 10662
QY 5463 TCGGCAATTTCAAGCATCTCTGCTGTATGCTACAGCATAGTACCGGACAGTAC 5522
DB 10663 TTGAGACATTTCAAGCTATCTCTTGAAGAGATCTTCACTGCGCAGACAGACTTA 10722
QY 5523 GCGTGTGAAGCTTCTGTCAAGAACATCCAGCTCCCTACCCCAAGCATATAGAGGT 5582
DB 10723 GGTATCTGAAGCTTCCGCGCAAGACGTGATGCTCCGTACAGGCGCGCATAGAT 10782
QY 5583 ATGAATGTGAAGAACACTCAAGACGACCCCTGCAAGAACACCAATTTGATGTA 5642
DB 10783 TTGAGATGTGAAGAACCACTCAAGCGCGCCATGCAAGAAACCGCACTTTGGGGTGA 10842
QY 5643 AAATTAAGTGAAGCTCTGCGAGCGCTTATCTGTGTTTACGGGCAATCCCTATCTG 5702
DB 10843 AGATTAAGTGAATTCGCTCGAGCGGTGACGTGTATACGGGAACATTTCCATTTCT 10902
QY 5703 TTGACATCCCTGATGAGCTTTTGTGAGATCATCAGATCAACCAATTTTGAAGTTA 5762
DB 10903 TTGACATCCCGAACGCTGCTTTTATCAGACATCAGATGACCACTGTGTCTCAACATGA 10962
QY 5763 GCTGACAGTGAAGACTGATTTATTTGACAGACTTTGTGATTTCTTAACATTAAGT 5822
DB 10963 AATGTAAGTCAAGTGAAGTCACTTATTCAGCAACTTCGCGGAGATGACCACTGAGT 11022
QY 5823 ACAAGCTGACAGGAGGACATTTGTCAGTTTCACTTCCCTCAACGACGCTGTTTGA 5882
DB 11023 ATGTATCCAGCCCGAAGGTCAATGCCCGTATATTCGCAATTTGAGCAGCAACTTCC 11082
QY 5883 AGGAAGCAGCACACATGTGACTGCGTAGGAGCACTTAACATTAATTTAGACATGCA 5942
DB 11083 AAGATGCAAGTACATGTCTTGAAGAAAGAGACGGTGAACATTTAGACCGCGA 11142
QY 5943 GCCCACAAGCAAAATTTATATGTTTGTGATGCGCAAGAAATGCACCTGCAATGCTGAAT 6002
DB 11143 GTCCACAGGGAATTTATGATGCTGTGTGTGGAAGAAACAACTGCAATGACGAAT 11202
QY 6003 GTAAACCAACGCGCGACCATATATTTGAGAACCAATTAAGTGCACAAAGATTTCCAG 6062
DB 11203 GTAAACCAACGCTGACATATGTGAGCAACCCCGACAAAAATGACCAAGAAATTTCAAG 11262
QY 6063 CGGAGTTTCAAAACATCTTGAAGTGTGCTGCTGCACTGTTTGGGGAGCATCATCC 6122

Db 11263 CCGCATCTCAAAAACATCATGAGTGGCTGCTGCTTTCCGCGCGCTGCTGC 11322
Qy 6123 TCATGTTGTAGGACTTATAGTGTGCTGTGAGCTCATGCTTAAACACAGTACT 6182
Db 11333 TATTATATATAGGACTTATAGTGTGCTGTGAGCTCATGCTTAAACACAGTACT 11382
Qy 6183 GACTGAGCGGACGACACTGA-CATAGCGGTAAACTGATGACTTCCGAGAAAGCTGCT 6241
Db 11383 GACCGCTACGCGCCCAATATATATGCTGACGAGAAACTGATGACTTCCGAGAAAGCTGATGT 11442
Qy 6242 GCATATGCGACGCGCGCT 6261
Db 11443 GCATATGCTACGAGCTGCT 11462

RESULT 10
US-09-102-248-8
Sequence 8, Application US/09102248
Patent No. 6008035
GENERAL INFORMATION:
APPLICANT: Johnston, Robert E.
APPLICANT: Davis, Nancy L.
APPLICANT: Simpson, Dennis A.
TITLE OF INVENTION: System for the in vivo delivery and
TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Seitzer Park & Gibson, P.A.
STREET: 1211 East Morehead Street
CITY: Charlotte
STATE: No. 6008035th Carolina
COUNTRY: USA
ZIP: 28234

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/102,248
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/801,263
FILING DATE: 19-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-147
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 11703 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-102-248-8

Query Match 36.8%; Score 2388.4; DB 3; Length 11703;
Best Local Similarity 64.7%; Pred. No. 0;
Matches 3698; Conservative 0; Mismatches 1981; Indels 41; Gaps 8;

Qy 580 ACTGACGATGAAACCGGAGCTATATTTCTCATCGAAACAGCCAGGTCACCTTC 639
Db 5746 ACTGACGATGAAACCGGAGCTATATTTCTCATCGAAACAGCCAGGTCACCTTC 5805
Qy 640 AACGAAATTCAGTACGTAATGTAATCAACAAAGAACTATATTTGATGGGCGCTCATG 699
Db 5806 AAAAGAGTCCGTTCTGAGAAACCAAGCTTACAGAACCAAGCTTGGAGCGCAATGCTCGG 5865

Qy 700 AGAATATTACGCCCCCGGCTCGATCTCGAAAGAGAAATGTTACAGAAAGTGC 759
Db 5866 AAAAGATTCATGCCCCCGGCTCGATCTCGAAAGAGAAATGTTACAGAAAGTGC 5925
Qy 760 AATTATGCGCTCGAAAGAAATGAAAGCAGTATCAATCAAGAAAGTAAATATGA 819
Db 5926 AGATGATGCCACCGAAGCCAAAGAAAGTATGATGACGATCTGTAAGTAAATATCA 5985
Qy 820 AAGCAATTCAGCGGAGGATCTATTTCTGATTTGGGCACTATCTATCATCAGAAATGA 879
Db 5986 AAGCAATTCAGCGGAGGATCTATTTCTGATTTGGGCACTATCTATCATCAGAAATGA 6042
Qy 880 ATCTGTGAGTGTATACAGATGATCAATTATCTGTACCAATCTACTGTCGCAAGGTAATTA 939
Db 6043 ATCAAGCAGATGCTATTAAGATCACTATCTGAAACCAATTTACTCTCGATGCTACCGG 6102
Qy 940 ACAGGTTTACATCTGCAAGAGTCCGCTTAAACGTCACCTATGTTATCCAGAGAAAT 999
Db 6103 CGAATCTCTCGATCCAGATTCGCTGATGCTGTGTACCAATCTATCTGATGAGAACT 6162
Qy 1000 ACCCTACAGTACGCAATTTATGTAATACAGATGAAATGATGCTATCTTACATGCTG 1059
Db 6163 ATCCGACAGTACATCTTATCAGATTACTGACGATGATGATGCTTACTTGTATGCTAG 6222
Qy 1060 ACGGCGATGCTGCTGTATGATACAGCACTTTTGGCGGCTAAACTGAAAGCTAC 1119
Db 6223 ACGGCGATGCTGCTGTATGATACAGCACTTTTGGCGGCTAAACTGAAAGCTAC 6282
Qy 1120 CAAGAAAGCATAGCTATTTTGCAGCCAGAGATGATGATGATGATGATGATGATGATGAT 1179
Db 6283 CAAGAAAGCATAGCTATTTTGCAGCCAGAGATGATGATGATGATGATGATGATGATGAT 6342
Qy 1180 ATCAATTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1239
Db 6343 ACAGCTACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6402
Qy 1240 GAGAAATACCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1299
Db 6403 GTGAACTGCAACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6462
Qy 1300 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1359
Db 6463 GTAAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6522
Qy 1360 TTACGCAATATGTAACAAAGCTGAAAGGCGGAAAGCAGCAATGTTTCCGAATCTC 1419
Db 6523 TCACCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6582
Qy 1420 ATATATCAAAACCGTTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1479
Db 6583 ATAAATTTGCTCCATTCAGAAAGTCCCTATGATGATGATGATGATGATGATGATGATGAT 6642
Qy 1480 ATGTCAAAAGTATCTCCGCGACGAAACATACAGAGAGCGGCTTAAGTGCAGGTTATTC 1539
Db 6643 ACGTGAAGATTAACACAGGACGAAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6702
Qy 1540 AGGCTGAGATCCCTGCTACGCTTACCTTTGCGGATGATGATGATGATGATGATGATGAT 1599
Db 6703 AAGCGGAGAAACCCCTGCGACTGCTTATGATGATGATGATGATGATGATGATGATGAT 6762
Qy 1600 GACTGATGCGGCTCTCTGCAATATCAATCTCTCTGATGATGATGATGATGATGATGATGAT 1659
Db 6763 GCTTACGCGCGCT 6822
Qy 1660 TTGATGCAATATGTAACAAATTTCCACACGCGGACCCAGTATTTGGAACGAGACATCG 1719
Db 6823 TTGATGCAATATGTAACAAATTTCCACACGCGGACCCAGTATTTGGAACGAGACATCG 6882
Qy 1720 GCTGCTTGAATTAAGGAG 1779
Db 6883 CATCATTCGACAAAG 6942

1780 ACTTAGGTGACCAACCGCTTGAATTTGATAGAGCGCGCTTGGCAATATACAT 1839
1843 ACCGGGTGTGATCAACCACTACTCGACTGATCGATGCGCTTTGGAGAAATATCAT 1902
1840 CTGTGACACTTACCTACAGAAAGAGTTTAAATTTGGTCCATGATGAAATCCGGTATGT 1899
7003 CCACCCATCTACCTACGGGTACTCGTTTAAATTCGGGGCATGATGAAATCCGGAATGT 7062
1900 TCTTAAAGCTGTGTGTGACACACTAGTCAATATCATGATTTGTACGAGAGTACTAGGTG 1959
7063 TCCCTACACTTTTGTTCMAACAGATTTGAAATGTCTATGCGCAGAGAGTACTTAAAG 7122
1960 AACGGTTAACCAAGCTGACGCGCGGCTCTATCGCGAGATTAACATAGTGCATG 2019
7123 AGCGGCTTAAACCTCCAGATGTGTGACGCTTCTTTGGCAGCAACATCTCATATGAG 7182
2020 TCGTCTCGACACCTTGAATGCGAGAGATGCGCCTTGGCTGAACATGGAATGAAAA 2079
7183 TAGTATCTGACAAAGAAATGCTGAGAGGTGCGCCTGCTCAACATGAGGTTAAGA 7242
2080 TTATTTGATGAGTATTTGATTCAAAGACCTTCTGTGGGGATTTATCCTGGTGG 2139
7243 TCATCGACGAGTATCGGTGAGAGACCACTTCTGCGGGATTTATCTTGGCAG 7302
2140 ACCAGATPACAGGACAGCTGACAGAGTGCAGACCTCTTAAAAAGCTTTTAAAGCTTG 2199
7303 ATTGGTTACTTCCACAGCGTTCGCGCGTGGCGAGCCCTTAAAAAGGCTGTTTAAAGTTGG 7362
2200 GAAAAACATTTGCCAGTGTGATATACCAAGCTGCGACCGCGCGGCACTGATGATG 2259
7363 GTAAACCGCTCCAGCGCAGACGAGCAAGAAAGCAAGAAAGCGCCTCTGCTAGATG 7422
2260 AAGCAATCCGATGGAACAGAAATTTGGAATTAAGGAGATTTAGAGGCGCTAGAAATCCA 2319
7423 AAAAAGAGCGGTGTTTAAAGATGATTAACAGGCACTTTAGCAGTGGCTGTGACGACC 7482
2320 GATACGAGATCATCTGCGACGCGCTGATCATCGCTCTGTCCACGTTTACCGGAAAGCG 2379
7483 GGTATGAGTGAACAATTTACACTGTCTACTGTGCACTTGAAGACTTTTGGCCAGAGCA 7542
2380 TTAAGAACTTCAAGAGCATTAAGAGGAGCCCAATCAACCTCTACGCTGACCTTAATAG 2439
7543 AAAGGACATTCAGAGCATCAAGAGGGAATAAAGCATCTCTACGCTGTCTTAATAGT 7602
2440 TGACGTAGTACAGACGACCTTAC-----CCACCGGAGAAATGTTTCAATACC 2466
7603 CAGATATGATATTCTACTACTAATTAATACTAACACACACACATGATAGAGGATTTCT 7662
2487 CTCAGCTGAACCTTTCACACGATTTTACCTTACAAATCCGATGAGCTTACCGAGATCAAA 2546
7663 TTTAATATCTGGCGCGCGCGCTTCCGCGCCCGCCACCTGCTCATGTGAGGCGCGGAGAA 7722
2547 CTCCTAGCGCGCGCTGAGGCGGCTTTCGAGCCCGCGCTGCTCAAAATCGAAGATCTTA 2606
7723 GAGGAGAGCGGCGCGCGGCTGCTGCGCAAGGAGGCTGCTTCAAAATTCAGCAACTGA 7782
2607 GAGAGTCAATAGTCACTTGAATTTCAAAACAACATCACTTAATCCGCGCGAGTCCA- 2665
7783 CCAAGCGGTAGTCCCTAGTCAATTTGAGACGAGCACTTAAGCTTCAACCCCGACGTCCAC 7842
2666 -----CCGCGCAAGAAAGAAAGAGTGTCTTAAAGCCAAACCTTACGCTTAAAAAGA 2720
7843 GCCCGCCACCGCGCGAGAAAGAGGCGCGCCAGCAACACCGAAAGCGAAAGAAACCA 7902
2721 AGAAGCAGCAAGCCAAAGAGAGCAAAAC---GCAAGCTTAAACAGGAAACGACACGTA 2777
7903 AAACGAGAGAAAGAAAGAAAGCAACTGCAAAAACCAAAACCGGAAAGAGACAGCGCA 7962
2778 TGTGTATGAAGTTGGAAGTGCAGCAAGACATTTCCGATCATG---CTGAACGCGCAAGTGA 2834
7963 TGGCACTTAAGTTGAGGCGCAGACATTTGTTGACGTCAAGAAACGAGGACGAGATGTCA 8022
2835 ATGATATGCTGCTGCTGCGAGAGAGGCTGATGAACAACCTCAAGTTGAAGAAAA 2894

8023 TCGGAGACGCACTGGCCATGAGAGAAAGGTATGAAACCTCTGACGTAAGAAAGCA 8082
2895 TTGATTAATGACAAATTAACGCGCGCTGAATTTGAAGAAAGGTACATGATACATTTGAGAT 2954
8083 TCGACCAACCTGTGTATCAAAAGCTCAAAATTTTACAAAGTGTGACATACGATGAGAT 8142
2955 ACGCGAGCTTCCCGAAGACATGAATTCAGACACGCTGCGAGTACACAGCGCAAAACAC 3014
8143 TCGCAAGTTTCCAGATCAACATAGAAAGTGAAGCATTTACCTTACACAGTGAACACCCCG 8202
3015 CGGCGCTTCTAACAATGACACACAGCGCGAGTCCAGATGAGAAATGGAATTTTACCGTAC 3074
8203 AAGGATTTCTAATCTGGCAACACAGAGCGGTGACAGTATAGTGAAGTATTTACCATCC 8262
3075 CGAAGAGAGTGGCGGGGAAAGCGACAGCGGAAAGACCGATCTCTGACACACAGAGCGAG 3134
8263 CTGCGGAGTGAAGAGGAGAGAGACAGCGGTGTCCGATCATGATTAATCTCGGTGCGG 8322
3135 TTGTGGCTATTTCTTAAAGAGGTGCAAAATGAGGGAAGCGGTACGCGGCTTTCAGTGTCA 3194
8323 TTGTGCGATAGTCTCTGAGTGAAGCTGATGAGAAACAGAACTGCCCTTTGCGTGTCA 8382
3195 CTGGAACCAAGAAAGGAGTACCATTAAGGATACCCCGAAGGTTTGAACCGTGT--- 3251
8383 CTGGAATAGTAAAGGAAAGACAAATTAAGACAGCCCGGAAAGGAGCAAGAGATGTCG 8442
3252 -----CACTAGTTACAGCGCTATGCGTCTTTGCAATGTCAAGTTCCTTGGCAAAAC 3305
8443 CAGCACACCTGAGTACGCGCAATGTGTTGCTCGAAATGTGACCTTCCATGAGACCGCC 8502
3306 CACCGGTGCTATTAACCTGACCGCAAGCAACCTGACGCTGCAAGAGAAACGTCG 3365
8503 CGCCCATCTATACCTCCGGAACCTTCCAGAGCCTTCAACATCTTTGAAGAAACCTGA 8562
3366 ACAATCCAAATTAACGACAGCTGAGAGAACTCTTGAATGTCATCAC---GCCGAC 3422
8563 ACCATGAGGCTTACGATACCTGCTCAATGCAATTTGCGGTGCGGATCGTTCGCGAGAA 8622
3423 CCAAGCAAGATTAACGATGACTTCACTGACAGTCCCTTACCTGAGGTTCTGCGCT 3482
8623 GCAAAAGAACGTCACAGACCTTAACCTGACAGCCCTTCTTGGGACATGCTCGT 8682
3483 ATTGACACACTAAGCGCGTGTTCAGCGCCCAATTAATAATGGAACGTTGGGACGAAT 3542
8683 ACTGCCACCTATGAAACGCTGCTTACGCTTGAATGACAGAGTGTGGAGCAAG 8742
3543 CTGATGATGATGATTAAGATCCAGTCTCGGACAAATTCGCTCAATCAGGCAAGCA 3602
8743 CGAGCATTAACCATACGATACAGACTTCCGCGCAGTTTGGATACGACCAAGCGGAG 8802
3603 CTGCGGATGACCAAAATTCGTTTACATGTCTTTTCAACAGACGATGATCAAGAG 3662
8803 CAGCAAGCGCAAAACAGTACCGCTACATGTGCTTGGCAGGATCAACACGTTAAAGAA 8862
3663 AAGATGAGAAATGACTATGACGACATCTGGAACCTGCGCTGCTTGGCACAAG 3722
8863 GCACATGATGATCAATTAAGATTAAGCCTTCAAGAGCGGTATGAAGCTTACGTAAG 8922
3723 GGTATCTCTGTAGTCAATGTCTCTCAGAGTGAACAGTGTAACTGATACAGAGCG 3782
8923 GATATCTTCTCTCGCAAAATGCGCTTCAAGGAGCAGCGTAAAGGTTAGATAGTGA 8982
3783 GAGCATTTGAGATTAATGACAGCTGTGAGAAAGAAAGATCAGAGAGAAATTTGTGCTGAG 3842
8983 GCAACTGAGCAAGTCAATGTAACACTGCGCGCAAGATTAACCAAAATTTGTGAGAGCGG 9042
3843 AGAGTACTTGTTCACCGCTCATGAGAAAGTGTGAAGTGCACGTTTACGATCAT 3902
9043 AAAAATATGATCTACCTCCGCTTACGCTTAAAAATTTCTTGCACAGTGTAGACCGCTC 9102
3903 TGAAGAGAGCTGCGCGGTATACATACATGACAGGCGCACGCGTATAGT 3962

Db 9103 TGAAGAAACAAGCTGAGCTACATCACTATGCAAGGCGGAGCCGCTTATACAT 9162
Qy 3963 CCTATCTGAGAGAGGCTCAGGCGAAGTGTATTAACACCTTCTGGCAAGACGTCA 4022
Db 9163 CCTACTGGAAGAAATCATCATGAGAAAGTTACGCAAGCGCGCATCTGGAAAGAACATTA 9222
Qy 4023 CCTACGAATGTATAGTGGCGACTACAGACAGTATCGTGAACGCGCAAGCAAGATGA 4082
Db 9223 CATTAGAGTGAAGTGGCGGACTACAGACCGGAACCGTTTCGACCGCGACCGAAATCA 9282
Qy 4083 AGCGTGCCTTAAGCAAAACAGTATGCTTCAAGAGCGCAACCAAGAAATGGGTCT 4142
Db 9283 CTGTTGTGACCGCCATCAAGAGTGTGTCTCTATTAAGCGCAACCAAGAAATGGGTCT 9342
Qy 4143 TCAACTCGCGGATCTTATTAGGCAACAGACCACTCAGTGAAGTAAATGGCAATTTC 4202
Db 9343 TCAACTCAGCGGATCTTATTAGGCAACAGACCACTCAGTGAAGTAAATGGCAATTTC 9402
Qy 4203 CATTCCGCTTGAACCGACAGTCTGCGCGTTCCTGTAGCTCAACGCGCTTACAGTCA 4262
Db 9403 CTTTCAAGTTGATCCCGAGTACCTGATGTCTCCCTGTTCGCAACGCGCGCAATGTAATC 9462
Qy 4263 AGTGTTCAAAGGCAATCACTCTTCACTGATGCAATGCGCAACATTGCTGACAAACA 4322
Db 9463 ATGGCTTTAAACACATCACTCCTTCAATTGATACAGACCACTTGAATTGCTCACACA 9522
Qy 4323 GAAATTTGGGGCTGGAGACGACGCAACGAGAAATGATTAAGGTTTATATCCAGA 4382
Db 9523 GGAAGCTAGGGGCAAAACCGGAACCAACATGAAATGATCGTGGAAAGAGGTCAAGA 9582
Qy 4383 ATTTTCTGTGGGGGAGAGAGGCTGAGTACGATGAGGTAAACATGAAACAGTCAAG 4442
Db 9583 ACTTACCGTGCACCGAGATGCGCTGGAATACATATGGGAATCATGAGCCAGTGAAGG 9642
Qy 4443 TCTGGGCCCCAGAGTGGGCAACAGGCAACCAATGATGCGCGATGATCATCTCC 4502
Db 9643 TCTATGCCCCAAGTACAGCAGCAGGACCTTCAAGATGCGCACAGAAATGATACAGC 9702
Qy 4503 ACTATTATCATGGGATCCAGTCTACATGTCATGTGTGTGTGTGTGTGTGTGTGTGT 4562
Db 9703 ATTACTACCATGCGATCTGTGTACACATCTTACCGGTGCGATCAGTACCGTGGCA 9762
Qy 4563 TCTCTGAGGCACTGATCATCAGAGCTTGGATGCGCAAAAGCAAGAGAGCTGCTGA 4622
Db 9763 TGATATTGGCGTAAACCGTTGAGTGTATGTGCTGTAAACCGCGCTGAGTGTCTGA 9822
Qy 4623 CGCCATAGCGCTTGCACCGAAACGCAACGGAACCAACAGATTAAGCGGTTTGTCTGA 4682
Db 9823 CGCCATAGCGCTTGCACCGAAACGCGTAACTCCAACTTGTGCGACCTTGTGTCTGG 9882
Qy 4683 TTGGGCAACCAACGCTGAAATTTGGAGAACTTTGAACATCTGTGTGTTAAACACC 4742
Db 9883 TTAGTTCGGCCATGCTGAAAGTTTCAACGAAACATGATTACTTGTGTGAAACAGTC 9942
Qy 4743 AACCTTTCTCTGGGCAAGTGTGATTCCTCTGGAGCGCTGTGTAATTCCTGCTCGT 4802
Db 9943 AGCCCTTCTCTGGGCAAGTGTGATTCCTCTGGAGCGCTGTGTAATTCCTGCTCGT 10002
Qy 4803 GCTTTTCATGCTGATGCTTTTATTTATGTTGACAGCGCTGCTCTGGGGAAGTGAACG 4862
Db 10003 GCTGCTCTGCTGCTGCTGCTTTTATGTTGTGCGGCGCTTACCTGCGGAGATGACG 10062
Qy 4863 CCTTGAACATGCGCAACCTGTGCAAAATGTTCCGGGATTCCTGTTAAAGGCTGTGCG 4922
Db 10063 CTTAAGAACATGCGCAACCTGTGCAAAATGTTCCAGATTAACGTAAGGCACTTGTG 10122
Qy 4923 AACGCGAGGTATACCGCACTTAACTGAGATCAAGGCTCTCATCGAATTAACAC 4982
Db 10123 AAAGGCGAGGTATACCGCGCTCAATTTGAGATCACTGTATGTCCTCGAGGTTTGC 10182
Qy 4983 CTTCAACTTAACAGAGTACGTACCTGCAATTCACACAGTCAATTCCTTACCAACAG 5042
Db 10183 CTTCAACCAACCAAGTACATTAACCTGCAATTCACACAGTGTGTGCTCCCTCCCAAAA 10242

Qy 5043 TTAATGCTGGGGGCTCCCTCGAGTGAAGGCACTCTCAAGGCGGATTAACATGCGCG 5102
Db 10243 TCAAAATGCTGGGGCTCCCTCGAGTGAAGGCACTCTCAAGGCGGATTAACATGCGCG 10302
Qy 5103 TTTTGGGCGTGTGATCCCTTTCATGTGGGAGCGCAATGCTTCTGTGACAGTGA 5162
Db 10303 TCTTGGAGGGGTCTAACCCCTTTATGTGGGAGGCGCAATGTTTTCGACAGTGA 10362
Qy 5163 ACACAACTGATGATGAGCGCTACGTGAGTTCGCTTCCAGACTGCACTTAATACGCA 5222
Db 10363 ACAGCAATGATGATGAGCGCTACGTGAGTTCGCTTCCAGACTGCACTTAATACGCA 10422
Qy 5223 TCGACTTAAAGTTTCAACAGCTGCTGAAAGTGGCGCTGTATTAATTAACGCAACA 5282
Db 10423 AGCCATTAAGGTGCACTGCTCGAGTAAAGTGAAGTGTGATTAATTAACGCAACA 10482
Qy 5283 CCACCGCGACCTGATACGTTTGTCAATGCGCTACCGCGAGTTCCTCAGGGAAGTGA 5342
Db 10483 CTACAGTTTCTGATGTGTATGATGAGAGAGTCAACAGGAACGCTTAAGACTTGA 10542
Qy 5343 AGGTATGACAGGCGCATATCAGCGCTTTTTCACCTTTGACCATTAAGTGTGATCA 5402
Db 10543 AAGTATGCTGACCAATTTACAGCATGTTTACGCGCATTTGATCATTAAGTGTGATCC 10602
Qy 5403 GAAAGGGGCTTTTAAACATACAGCTTCCCTGAGTATGAGGCTTAAGCAAGAGGT 5462
Db 10603 ATCCGGGCTGTGTACATCATGATCTTCCGAAATATGAGCGATGAACCGAGAGGT 10662
Qy 5463 TCGGCAATTAATCAAGCATCTCTGCTTGAATGCTACAGATGATGAGCGCGCATATC 5522
Db 10663 TTGAGACATTAATCAAGCATCTCTGCTTGAATGCTACAGATGATGAGCGCGCATATC 10722
Qy 5523 GCGTGTGAAGCTTCTGTCAAGAAATCCAGTCCCTTACACCCAGACGATATCAGGT 5582
Db 10723 GCGTGTGAAGCTTCTGTCAAGAAATCCAGTCCCTTACACCCAGACGATATCAGGT 10782
Qy 5583 ATGAATGTGAAGAAACATCAGGAGAGCGCGCGCAAGAAACAGACATTTGATGATGA 5642
Db 10783 TTGAGATGTGAAGAAACATCAGGAGAGCGCGCGCAAGAAACAGACATTTGATGATGA 10842
Qy 5643 AAATTAAGTGAAGCTTCTGTGAGAGTCTAATCTGTGCTTACGGGCAATCTCTGA 5702
Db 10843 AGATTGCAATTAATCTGCTCGAGAGGAGTGTATGAGGAACATTTCCATTTCTA 10902
Qy 5703 TTGACATCTCTGATGAGCTTTTGTGATATCAAGAAATCAACCAATTTTGAAGTTA 5762
Db 10903 TTGACATCTCTGATGAGCTTTTGTGATATCAAGAAATCAACCAATTTTGAAGTTA 10962
Qy 5763 GCTGACAGTGAAGACATGATATTTATCTGAGACTTGTGTGTCTTAACTATGACGT 5822
Db 10963 AATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 11022
Qy 5823 ACAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 5882
Db 11023 ATGATTCGAGACCGGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 11082
Qy 5883 AGAAGCGACCAACATGATGATGCTGAGGAGACATTAACATTAATTAAGCAATGCA 5942
Db 11083 AAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 11142
Qy 5943 GCCCAAGCAAAATTTATTAATTTGCTGATGCGGCAAGAAATCCACCTGCAATGTAAT 6002
Db 11143 GTTCAAGCGCAAACTTATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 11202
Qy 6003 GTTAAACCAACGCGGAGCAATTAATTTGAGAAACCAATTAAGTGAAGCAAAATTCAG 6062
Db 11203 GTTAAACCAACGCGGAGCAATTAATTTGAGAAACCAATTAAGTGAAGCAAAATTCAG 11262
Qy 6063 CGGAGTTTCAAAACATCTTGAAGTGTGCTGCTTGAACGTTTGGGGAGCATATCC 6122
Db 11263 CGGAGTTTCAAAACATCTTGAAGTGTGCTGCTTGAACGTTTGGGGAGCATATCC 11322

OY	1840	CTGAGCACTTACCTTAACAGAAAGAGTTTAATTTGGCCATATGAAATCCGGTATGCT	1839
Db	7003	CCACCAATCTACCTTAAGGGTACTCGTTTAAATTCGGGGCCATATGAAATCCGGAAATGT	7062
OY	1900	TCTTAACCTCTTTGTCAACACACTAGTCAATATCATATGTTGTAAGCAGATACACGTG	1959
Db	7063	TCCTCACACTTTTGTCAACACAGTTTGAATGTGTATTATGCGACGAGTACTAGAG	7122
OY	1960	AACGGTTAACACCGTCAGCGGTGCGGGCCCTCTATCGCGGAGATTAACATAGTCATG	2019
Db	7123	AGCGGCTTAAACGTCAGATAGTGCAGGGTTCATTGGGAGCGACAACTCATATCATGGAG	7182
OY	2020	TCGTCTCCGACACCTTGATGCGCGAGAGATGCGGCACCTTGGCTGAACATGGAATAAAA	2079
Db	7183	TAGTATCTGACAAAGAAATGGCTGAGAGTGCGCCACTGTGCTCAACATGAGAGTTAAGA	7242
OY	2080	TTATTTGATGCAAGTTATTTGGTATCAAAGACCCCTACTGTGGGGGATTTATCCGTGGTG	2139
Db	7243	TCATCGACGCACTCATCGGTAGAGACCACTTACTTTGCGGGGGAATTTATCTTGCAAG	7302
OY	2140	ACCAATTAACAGGCAACAGCTGCGAGAGTCGAGACCTCTAATAAAGCTTTTAAAGCTTG	2199
Db	7303	ATTGGGTACTTCCACACAGCGTGGCGCGTGGGGAGCCCTGATAAAGCTGTTAAAGTTGG	7362
OY	2200	GAAAAACCTTGGCAGCTGCATGATATACCAAGACTGGACCGCCCGGGGCACTGCATGATG	2259
Db	7363	GTTAAACCCCTCCAGCCGACGACGAGCAGAAAGCAGAAAGACGCGCTCTGCTTAATG	7422
OY	2260	AAGCAATGCAGTGAAGAAAGAAATTTGGAATTAGCGACGAATTTGTGAAGGCGCTTAATCCA	2319
Db	7423	AAACAAAGGCGTGGTTTGAAGTGAATTAACAGGACCTTTAGCAGTGGCCGTGACGACC	7482
OY	2320	GATACGAGATCATACTGGCAGGCGCTGATCATACGTCCTGTCCACGCTTACCCAAAGCG	2379
Db	7483	GGTATGAGGTAGACAAATTATACACCTGTCTTACGTGCAATTAGAACTTTTGCCAGAGCA	7542
OY	2380	TTTAAGAACTTCAAGAGCATTAAGAGGAGGCCCAATCACCTCTAGCGCTGACCTTAATAGG	2439
Db	7543	AAAGAGCATTCCTCAAGCCATCAGAGGGGAAATTAAGCATCTCTACGGTGTCTTAATAGT	7602
OY	2440	TGACGTAGTAAACACGCACTTAC-----CTACCGGCGAATGTTTTCATATCC	2486
Db	7603	CAGCATATGATCATTTGATCTGACTAATATCTAACACACACACACATGATTAAGAAATTC	7662
OY	2487	CTCAGCTAACTTTTCCACACAGTTTAACCTTACAAATTCGATGCGTTCACGAGATCCAAACC	2546
Db	7663	TTTAACATGCTGGCGCGCCCTTCCCGGCCCTCCACCTGCATGTGAGGCGCGGAGAA	7722
OY	2547	CTCTTAGGCGCGCTGAGAGCGCTTTTGGCCCCCGCTGGCTGCTCAATTCGAAGATCTTA	2606
Db	7723	GGAGGACAGGCGGCCCGCATGCTCGCGGCAACGGGCTGGCTTCTCAATTCAGCAACTGA	7782
OY	2607	GGAGGTGATAGTCAACTTGACTTTCAAAACAAGTACCTTAATTCGCGCGCAGGTCCA-	2665
Db	7783	CCACAGCCGTCAGTGCCTTACTGATTTGACACGAGCACTAGACTCAACCTCCACAGTGCAC	7842
OY	2666	-----CCGCCAAAGAAAGAAAGAGTGTCTTAAGCCAAACCTAATCAGCTTAAANAGA	2720
Db	7843	GCCCGCCACCGCGCCAGAAAGAGCAGGGCGCCAAAGCAACACAGAGCCGAAGAACCA	7902
OY	2721	AGAAGCAGCAAGCCAAAGAGAGCAGAAAC--GCAAGCTTAAACAGAGGAACGAAACGTGA	2777
Db	7903	AAAGCAGAGAAAGAAAGAAAGAACCTCGAATAAACCAACCCGGAAGAGAGCAGCGCA	7962
OY	2778	TGTGTATGAAGTTGGAGTTCGACAAAGACATTTCCGATCATG--CTGAACGCGCAAGTGA	2834
Db	7963	TGGCACTTAAGTTGAGGCGCCAGATTTGTCAGCTCAAGAAAGAGAGCAGAGATGTCA	8022
OY	2835	ATGGAATAGCCTGGGTGTGCGAGGAGAGCGTGAAGAAACCATCCTCAGTTGAAGAAAA	2894
Db	8023	TGCGGACAGCACTGGCCATGAGAGGAAGTAATGAACCTCTGCACTGTGAAGAAACCA	8082

QY	2895	TTGATTAATGACCAATTAGCGGCGGTGAATTGAGAAGGCGTAAGCATGTACGACTTTGGAGT	2955
Db	8083	TCGACACACCTGTGCTATCAAAAGCTCAAAATTTCCAAAGTCTGTACGATACGACATGAGGT	8142
QY	2955	ACGGGACGTTCCCGCAGAACATGAAATAGACACGCTGCACTACACCGACAAACAC	30145
Db	8143	TCGCACAGTTCCCACTCAACATGAGAAAGTAGAGCAATTCACCTACACAGTGAACACCCG	82025
QY	3015	CGGGCTTTCACACTGCGCACCGCGCGAGTCCAGTATGAGAAATGGAGATTTACCGTAC	3074
Db	8203	AAGGATTTCTATACTGGCACACCGAGGGGTGCAAGTATATGAGAGTGAATTTACATCC	8265
QY	3075	CGAAGAGGTGGGGGGAAGGCGACACGCGGAACCGATCTCTGGAACAACAGAGCAGAG	3133
Db	8263	CTCGGGAGTAGAGGGCGAGAGGAAACAGCGGATCTCGATATGATTAATCTCCGATCGGG	8322
QY	3135	TTGTGCTATATGTTCTTAGAGAGTGCAAATAGAGGCGACGCTGACGCGCTTTCAGTGTCA	3194
Db	8323	TTGTGCGATAGTCTCTCGGTGAGACTGTATGAGAACACGAACTGTCCCTTTCGGTGTCA	8388
QY	3195	CTTGGAAACAGAAAAGGGGTGAACATTAGAGATACCCCGAAGGTTCTTGAACCGTGT---	3251
Db	8383	CCTGGAAATAGTAAGGGAAGACAATTAAAGACGACCCCGGAAGGGACAGAAAGTGTGTCG	8442
QY	3252	-----CACTAGTTACAGCGCTATGCGGTCTTTGGAATGTACGTTCCCATGCGCAAAAC	3305
Db	8443	CAGCACCACTGTGTACCGGCAATGTGTTGTCTGGAAATGTGAGGCTTCCATGCGCACCGCC	8502
QY	3306	CACCCGATGTGTAATTCACTGAACGCGCAGAACGAAACACTGACGTGTGGAAGAGAACGTC	3365
Db	8503	CGCCCACTGTGTATACCGCGAACCTTTCAGAGCCCTGTGACATCTTTGAAGAGAACGTCGA	8562
QY	3366	ACAATCCAAATTACGACACGCTGTGAGAAAGCTTTGGAATGTTCATCAC---GCCGCGC	3422
Db	8563	ACCATGAGGCTTAGATACCTGTGCTCAATGCAATTTGCGGTGGGAGTCTGTGCGACAA	8622
QY	3423	COAAACGAGCATTAACCGATGATCTTACACTGACCAAGTCCCTACTGGGGTTTGCCTGT	3482
Db	8623	GCAAAAGAGCGTCACTGACACGACTTTTACCTTGACACAGCCCTTACTTGGGCAACATGCTGT	8682
QY	3483	ATTGCGAGCACTCAACCGCGGTTTCAAGCCCAATTAATAATTGAGAACGTGTGGAGCAAT	3542
Db	8683	ACTGCACCACTATCGAAACCGTGTTCAGCCCTGTTAAGATGACACAGTCTGGAGCGAAG	8742
QY	3543	CTGATGATGATTCGATTAGAAATCCAGGCTTCTGGCACAATTTGGGCTTACATCAGCAGCA	3602
Db	8743	CGAGCGATTAACCATATAGCATATACAGACTTCCGCCAGTTTGATTCGACCAAAAGCGAG	8802
QY	3603	CTGCGGATGTCACCAAAATTCGTTACATGTCTTTGACACGACGACATGACATCAAGAG	3662
Db	8803	CAGCAAGGCGCAAAACAAGTACCGGTACATATGTGCTTGGACAGAGTACACACGTTTAAAGAG	8862
QY	3663	ACAGTATGAGAAATATAGCTATTCAGCATATGTGCAACCTGCGCTGTCTTGGCCACAAG	3722
Db	8863	GCACCATGATGACATCAAGATTATGCACTTCAGGACCGTGTGAAAGCTTATGCTAACAG	8922
QY	3723	GGTACTTCTCTTAGCTCAATATGTCTCTCAGGTGACAGTTAACGTCAGTATCAGAGCG	3782
Db	8923	GATACTTTCTCTCGCAAAATGCGCTTCCAGGGGACAGCGTTAACGTTATGTAAGTA	8982
QY	3783	GAGCATCTGAAATTCATGACACCGTGAAGAAAAGATCAGGAGGAAGTTGTTCGGTAGAG	3842
Db	8983	GCAACTCGACAAAGCTATGTACACTGGCCGCGCAAGATTAACCAAAATTCGTGGGACGGG	9042
QY	3843	AGGAGTACTTGTTCACACCCGTCCATGAGAAAGCTGTGTAAGTGGCAGCTTTATCGATCACT	3902
Db	9043	AAAAATATGATCTACCTCCCGTTACCGGTAAAAAAATTCCTTTCGACAGTGTACACCGTTC	9102
QY	3903	TGAAGAGAGCGTCTGCGGGGTACATAACATGACAGGCGCACGGCCCAACGCGTATAGT	3962
Db	9103	TGAAAGAAACAACCTGCAAGCTACATCACTATGACACAGGCCGGAACCGCAACGCTTATACAT	9162
QY	3963	CCTATCTTGAGGAAGCGTCAGGCGCAAGTGTACATTAACAACCTTCTTGGCAAGAACGTCA	4022

Db	9163	CTTACCTGGAGAAATCATCAGGGAAAGTTTACCGAAAGCCGCATCTGGGAAAGAACATT	9222
Qy	4023	CTTACGAATGTAAATGTGGCACTACAGCAGAGTATCGTAGCAGCCGAAAGAAATGA	4082
Db	9223	CGTATGATGCAGATGGCGGCACTAACAGACCGGAACCGTTTGACCCGCACCGAAATCA	9282
Qy	4083	ACGGCTGACTTAAAGCAAAACATGCGATTTCCTTCAAGAGCCGCAAGAAATGGGTCT	4142
Db	9283	CTGGTTGCACCGCCATCAACCAATGGGTGCCTTATAGCCGACCAAGAAAGTGGTCT	9342
Qy	4143	TCAACTGCGCGGATCTTATTAGGACAACAAGCACTCAGTGCAGAGGTAATTCACATTC	4202
Db	9343	TCAACTACCGGACTTATATCAGACATGACACACGCGCCAGGAAATTCATTTGC	9402
Qy	4203	CATTCCGCTTGACACCGACAGTGTGCCGGTTCGTTAGTCAACAGCCTACAGTACGA	4262
Db	9403	CTTTCAAGTTGATCCGAGTACCTGCATGTGTCCCTGTGGCCACGCGCGGAATGTAAATAC	9462
Qy	4263	AGTGGTTCAAAAGGATACCCCTCCACTGACTGCATTCGCAATCGCAATTTCTGCACAGA	4322
Db	9463	ATGGCTTTTAAACATATAGCCTTCCAATTATATACAGACCACTTGACATTTCTACCACCA	9522
Qy	4323	GAAATTTGGGGCTGCGAGCAGACGCAACACGCAATAGSATTTACGGGTCTACATCCAGA	4382
Db	9523	GGACACTAGGGGGCAAAACCCGAAACCAACCTAATATGATCGTGGAAAGACGTTAGAA	9582
Qy	4383	ATTTTCTGTGGGCGAGAAAGGCTGGAGTACCTATGGGGTAAACAATGAACAGTCAAG	4442
Db	9583	ACTTCACGTCGACCGAGATGGCTGGATATACATATGGGAAATCATAGACCGATGAGGG	9642
Qy	4443	TCGTGGCCCAAGATCGGCAACGAGCGAACCAATGATGGCCGACATGATCATCATCC	4502
Db	9643	TCATATGCCCAAGATCAGCACACGAGAACCTCTACGSAATGGCCACGAAATATGTAACAGC	9702
Qy	4503	ACTATATCATTCGSCATTCAGTCTACATCTCATATGGCTGTGTGGTATCGCTCTTGCTA	4562
Db	9703	ATTATCTACATTCGCACTCTGTGTACACATCTTACCGCTGCATCAGTACCGTGGCGA	9762
Qy	4563	TCCGTGATAGGCACTGCACTCATCAGACAGCTTGCACTCCGCAAGAAAGAAAGACCTGCCTGA	4622
Db	9763	TGATGATTTGGCGTAAACGTTGCACTGTATATGTGCCTGTAAAGGCGCCCTGAGTGCCTGA	9822
Qy	4623	CGCCATACGCGCTTGCAACGCAACGCAACGGATCCACAGCATTTAGCGGTTTTGTGTGCA	4682
Db	9823	CGCCATACGCGCTTGCGCCCAAAACGCGCTAATCCCACTTCGCTGCGACATCTGTGCTGG	9882
Qy	4683	TTGGGCCAACCAACGCTGAAACATTTGGAAACCTTGAACCATCTGTGGTTTAAACAACC	4742
Db	9883	TTAGTGTGGCCAAATGCTGAAACGTTACCGAGACATGATTTACTGTGTGTGGAACAGTCC	9942
Qy	4743	AACCGTTTCTCTGGGCAAGTGTGCAATCTCTGCGACGCGCTGTATTTCTGTTCCGCT	4802
Db	9943	AGCGCTTCTTGTGGGTCCAGTTGTGCACTACTTTGTGGCCGTTTCACTGTTCTTAATCCGCT	10002
Qy	4803	GCTTTTCACTGTGCATCCCTTTTATTTATGTGTCAGGCGTCTGCTGGGAAAGTACAGC	4862
Db	10003	GCTGCTCTCTGCTGCTCCTTTTATTTATGTGTTCGCGCGCTTACTGGGAAAGTACAGC	10062
Qy	4863	CCTTCGAAACATGCGACACATCTGTGCAAAATGTTCCGGGGAATCCCGTATTAAGCGTTGTGC	4922
Db	10063	CTTACGAAACATGCGACACATCTGTCCAAATGTGCAAGATACGATTAAGGCACTTGTGTC	10122
Qy	4923	AACGCGAGGTTTACGCGCACTTAACTGGAGATCAACGGTGCCTCATGGGAATTAAACAC	4982
Db	10123	AAAGGCGAGGTTATGCCCCGCTCAATTTGGAGATCACTGTGCATGTCTCTGGAGGTTTTC	10182
Qy	4983	CTTCAACTTAAACAAGATACGTGACTTGAATTTCCACACAGTCATTTCTTCAACCAAG	5042
Db	10183	CTTCCACCAACCAAGATATTAATTCGGAATTTCAACATGTGTGTCCTCTCCCAAAAA	10242
Qy	5043	TTAAATGCTGGGGTCCCTGAGTGCAGGCAATCTCAAAAGCGGATTAACATGCGCG	5102

Db	10243	TCAAATGCTGGGGCTCTCTTGGAATGTCACCGGGCGCTCATGCAACAATACTACCTGCCAAG	10302
Oy	5103	TTTTTGCGGTGTGTAACCTTTTCATGTGGGAGGGCGACATGCTTTCTGTGACAGTGA	5162
Db	10303	TCTTCGGAGGGGTCTACCCCTTTATGTGGGGAGAGCGCAATGTTTGTGCGACAGTGA	10362
Oy	5163	ACACACAACAGTAAGTGGGGGTACGTGCAATTTGGCTCCAGACTGCACTTAATGATCAAGG	5222
Db	10363	ACAGCCAGATGAATGAGGGGTAGTGAACCTGTCAAGCAATTTGCCGTGTGCCAGCGGC	10422
Oy	5223	TCGCATTAAGTATTCACAGAGCTGCTCTGAAAGTCGGCTCGCTAGTATATACGGCA	5282
Db	10423	AGGGGATTAAGTGACACTCGCCCGCATGAAGTAGAGACTGGGTATATGATACGGGAACA	10482
Oy	5283	CCACCCGCGACCTTGATACGTTTGTCAATGGCGTCACGCAGGTTCTCACGGACCTGA	5342
Db	10483	CTACACAGTTTCTAGATGTGTACGTGACGAGGTCAACCAAGAACGTTTAAAGACTTGA	10542
Oy	5343	AGGTCATAGAGGGCCGATATACGCCGCTTTTACACCTTTGACATATAAGTCGCATCA	5402
Db	10543	AAGTCATAGCTGAGCAACATTTACAGATCGTTTACGCAATTCATATAAAGTCGTTATCC	10602
Oy	5403	GAAAGGGGCTTGTTTTCAACTACGATTCCTCTGAGTAGAGACTTATGAACCAAGAGCGT	5462
Db	10603	ATCGCGGCTGCTGTACACTATAGATTTCCCGGAATATGACACGATGAACCAAGAGCGT	10662
Oy	5463	TCGGCGATTTTCAAGACATCTCGCTTGATGTACAGACATAGTAGCCGCACTGACATAC	5522
Db	10663	TTGGAGACATTTCAAGCTACCTCTTACTAGCAAGAGATCTCATGCCACGACAGACATTA	10722
Oy	5523	GGCTGCTGAAGCCTTTCTGTCCAGAAATCATCAAGTCCCTTACACCCAGAGATACAGGT	5582
Db	10723	GGCTACTCAAGCCTTTCCGCGAAGAGTGCATGTCCCGTACACGAGCGCGATACAGAT	10782
Oy	5583	ATGAATGTGGAAGAACAACTCAGACGACCCCTGCAAGAAACAGACCACTTTGAGATGA	5642
Db	10783	TTGAGATGTGGAAGAAACAACTACAGCGCCGCCACTGCAAGAAACCGCACTTTGGGTGA	10842
Oy	5643	AAATTGAAGTGAAGCTCTGCGAGCGTCTTAATCTGTGCTTACGGGACATCCCTATCTGA	5702
Db	10843	AGATTTGACATTAATTCGGCTCCGAGCGGTGGAATGTTTATACGGGAACATTTCCATTTCTA	10902
Oy	5703	TTGACATCTCTGATGCAAGCTTTTGTGAGATCATCAGAAATCACCAACAATTTTGAAGTTA	5762
Db	10903	TTGACATCTCCGAACGCTGCTTTATAGAGACATCAATGCAACACTGTGCTTCAACAGTCA	10962
Oy	5763	GCATCAGTAGAGAGACTGCAATTTATTTGCAAGACTTTGGTGGTTCTTAACTATACAGT	5822
Db	10963	AATGTGAAGTCAGTGATGCACTTTATTCAGCAGACTTCCGCGGGAATGGCAACCTGCACT	11022
Oy	5823	ACAAAGCTGACAGGAGGGAACATTTGTCAGTTTCACTCCCACTCCAGACAGCTGTTTTGA	5882
Db	11023	ATGATATCCGACCCGGAAGGTCAATGCCCCGTAATTCGATTCAGACACAGCAACTCTCC	11082
Oy	5883	AGGAAGGACCAACATGTGATCTGCCGTATGGCAGACTTAACATTAATTTTAGCAATGGA	5942
Db	11083	AAGGTGACAGTACATGTCTTGGAAGAAAGAGCGGTGACATTAACATTTTACACCGGA	11142
Oy	5943	GCCCCAAGCAAAATTTTATGTGTGCTATNGGGGCAAGAGTCCACCTGCAATGCTGAT	6002
Db	11143	GTCACACGGGAACTTTATGCTATGCTGTGTGGGAAGAGCAACATGTCAATGCAAGAT	11202
Oy	6003	GTAACCAACCGGCGGACCACTAATTTGGAAGAACCAATTAAGTGCACCAAGATTTCCAG	6062
Db	11203	GTAACCAACAGCTGACCAATGTGTAGCAACCCCGACAAATAATGACAAAGATTTTCAAG	11262
Oy	6063	CGGAGATTTCCAAAACATCTTGAACCTGCTGTTGCACTGTTTGGGGAGCATATCCC	6122
Db	11263	CCGCACTCTCAAAAACATCATGAGATGGCTGTGTTTGCCTTTTCCGCGCGCTGTGTCG	11322
Oy	6123	TCAATGTTGTAAGACTTATATAGTGTGGTGTGCAAGCTTAATGCTTATTAACAACGTATAG	6182
Db	11323	TATTAATTAATAGACTTATATGATTTTGTCTTGACAGCATGATGCTACTACACACAGAT	11382

QY 6183 GACTGAGCGGACACTGA-CATAGCGGTAAACTCGATCTTCCGAGNAGCGTGT 6241
DB 11383 GACCGCTACGCCCATGATGTCGACAGAAAACCTCGATCTTCCGAGNAGCGTGT 11442
QY 6242 GCATATGCGACGCGCGCT 6261
DB 11443 GCATATGCTATCGCGTGT 11462

RESULT 12

US-08-801-263A-4
; Sequence 4, Application US/08801263A
; Patent No. 5811407
; GENERAL INFORMATION:
; APPLICANT: Johnston, Robert E.
; APPLICANT: Davis, Nancy L.
; APPLICANT: Simpson, Dennis A.
; TITLE OF INVENTION: System for the In Vivo Delivery and
; TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
; STREET: 1211 East Morehead Street
; CITY: Charlotte
; STATE: No. 5811407th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,263A
; FILING DATE: 19-FEB-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-147
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-420-2200
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11717 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-801-263A-4

Query Match 36.8%; Score 2386.4; DB 1; Length 11717;
Best Local Similarity 64.5%; Pred. No. 0;
Matches 3745; Conservative 0; Mismatches 1991; Indels 74; Gaps 9;

QY 580 ACTGACGATGAGACGGGAGCGTATATTTTCTGATCGGAAACAGGCCAGCTTC 639
DB 5761 ACTGCTACCGGGGTAGGTGGTACATATTTTCACGAGACACAGCCCTGGGCACTTGC 5820
QY 640 AACAGAAATCAGTACGTCAATGTAACAAAGAACTATATTTGGATCGGGCCGTCATG 699
DB 5821 AAATGAGATCCGCTTTCGAGATACGCTTACGAAACCGACCTTGAAGCGCAATGTTCTGG 5880
QY 700 AGAAGATTAACGCCCCCGCGCTCGATCTCGAAAGAGAGAAAATGTTACAGAAAGAACTGC 759
DB 5881 AAAGATCTACGCCCCCGGTGCTCGACAGTCGAAAGAGAAAGCAAGCTCAAACTCAGATACC 5940
QY 760 AATTAATGCGCTCTGAAGAAATTAAGCAGTATCAATCAGAAAGTAAAGTAAATGA 819
DB 5941 AGATGATCCCAACGAGCCCAACAAAGCAGGTACAGCTTGAAGAAATGAAATCAGA 6000

QY 820 AAGCAATTACAGCGAGGAGCACTCATTTCTGATTTGGGACATATCTATCATCAGAAATGA 879
DB 6001 AAGCATTAACCACTAGAGAGCTGTTTCAGGGCTACGACTGATTAACCTGCGACAG--- 6057
QY 880 ATCTGTGAGTGTTCACAGAGTCAATTTATCTGTACCAATCTACTCTGTCACAGTAATTA 939
DB 6058 ATCAAGCAGAAATGCTATTAAGATCACTTACCCGAAACCACTGATTCAGCAGGTACCGG 6117
QY 940 ACAGTTTACATCTGACAGAGGTGCGGTTAAAGCTGAACTTATTCAGAGAAAT 999
DB 6118 GCAACTTACTTGACCCCAAGATTTGCTGATGCTTTTGAACAACATATCTGCAATGAAAT 6177
QY 1000 ACCCTACAGTACCGCAGTATTTGATATACAGATGATGATGATGATGATGATGATGATG 1059
DB 6178 ACCGAGGTGATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6237
QY 1060 ACGGCGATCGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1119
DB 6238 ACGGAGCAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6297
QY 1120 CAAGAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1179
DB 6298 GCAAGAAGCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6357
QY 1180 ATACATTAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1239
DB 6358 AAGCTTCAACAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6417
QY 1240 GAGATTAACCTGTCTTGAATTTGCGGCAATTTATGTTGATGATGATGATGATGATGAT 1299
DB 6418 GTGATTTCCCAACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6477
QY 1300 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1359
DB 6478 GTATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6537
QY 1360 TTACGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1419
DB 6538 TTACGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6597
QY 1420 ATATCTTAAACCGTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1479
DB 6598 ATATTTGCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6657
QY 1480 ATGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1539
DB 6658 ACGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6717
QY 1540 AGGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1599
DB 6718 AAGCGGAGAAACCTGCGAGACCGTTTACTGCTGCGGATGATGATGATGATGATGATGATG 6777
QY 1600 GACTGAATGCGGTGCTTCTGCAATATCATCTCTTCTGCAATGTCAGCGGAAGAT 1659
DB 6778 GCGTTACAGCGGTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6837
QY 1660 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1719
DB 6838 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6897
QY 1720 GGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1779
DB 6898 CTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6957
QY 1780 ACTTAAGTGTGACCAACCGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1839
DB 6958 ACTGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7017
QY 1840 CTGTCAGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1899
DB 7018 CCACCATCTGCGACGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7077
QY 1900 TCTTAACGCTGTTGTCAACACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1959

Db 7078 TCCGACGCTCTTGTCAACAGAGTTCGAAATGCTGATATCGCCAGCAGAGATTAATGGAGG 7137
Qy 1960 AACGGTTAACCAACCTCAGCCGCGCCGCTTATCGCGGAGAGATTAACATAGTGAATGTG 2019
Db 7138 AGCGGCTTTAAACCTTCACAAATGTGACGATTTATCGGAGCAGCAACATCATACAGAG 7197
Qy 2020 TCGCTCCGACACCTTGATGGCGGAGAGATCGCCACTTGGCTGAACATGGAGTAAAAA 2079
Db 7198 TAGATATCGAACAAAGAAATGCTGAGAGTGTGCCACTGGCTCAACATGGAGTTAAGA 7257
Qy 2080 TTATATGACGATTATTTGATTAACAAGCACTTCTGTGGGGGATTTATCTGTGTG 2139
Db 7258 TCATTGACGAGTATCGGCGAGAGACCGCTTACTTCTGGGGGATTCATCTTTGCAAG 7317
Qy 2140 ACCAGATTAACAGGACAGCCCTGCAAGATCGCAAGCCCTTAAAAAGCTTTTAAAGCTG 2199
Db 7318 ATTGCGTTACCTCCACAGCGGTGTGCGGGGACCCCTTGAAAAGCTGTTTAAAGTTGG 7377
Qy 2200 GAAAACCATTTGCGAGTATACCCAGACTGCGACCGCGCGGCACTGCGATG 2259
Db 7378 GTAAACCGCTCCAGCCGAGCAGACAGACAGAGACAGAAAGACGCGCTTGTCTAATG 7437
Qy 2260 AAGCAATGCGATGGAACAGAAATTGAAATTACGAGAGTTAGTAAGCCGTAGAAATCCA 2319
Db 7438 AAACAAAGCGTGGTTTGAAGTAGTATTAACAGACACTTAGCAGTGGCGGTGCAACTC 7497
Qy 2320 GATACGAGATCATATCGCAGCGCTGATCATCAGTCTGTCTCCACGTTAGCCGAAAGCG 2379
Db 7498 GGTATGAGTAGACAAACATCACCTGTCTGTGCGATGGAACCTTTGGCCAGAGCA 7557
Qy 2380 TTAAAGACTTCAAGAGCATTAAGAGGAGCCCAATCACCTCTAGCGCTGACCTAAATAG 2439
Db 7558 AAGAGAGATTTCAAGCCATCAGAGGGGAAATTAAGCATCTCTAAGGTGTCTTAAATAGT 7617
Qy 2440 TGACGTAGTA-----GACACGACCTTACCCACCGGAGAAATTTTCCATPACC 2486
Db 7618 CAGCATGACATTTCTACTGACTAATATACACAACACACCATGAAATGAGGATTCCT 7677
Qy 2487 CTCAGCTGAACCTTCCACACGATTTAACCTTACAAATCCGATGGCTTACCGAGATCCAAAC 2546
Db 7678 TTAAACATGCTCGCGCGCGCCCTTCCCGCCCTCCACCTGATGAGCGCGGAGAA 7737
Qy 2547 CTCCTAAGCGCGCGCTGAGGCGGCTTTCGCGCCCGCTGCTCAATCGAATCGAATCTTA 2606
Db 7738 GAGAGCAGGCGGCGCCGATGCTTCCGCGCAATGGGTGGCTTCCCAATTCAGCAACTGA 7797
Qy 2607 GAGAGTGAATGATCACTTGACTTTCAAAACA-----ACGATCACTTAATCCGCGCGAG 2660
Db 7798 CCAACAGCCGTACGTGCTTACTGATGGAACAGGCACTAGACCTCAAAACCCACGCCAC 7857
Qy 2661 GTCCACCGCCAAAGAAAGAAAGAGTGTCTTAAGCCCAAACTACTACGCTTAAAAAGA 2720
Db 7858 GCCCGCGCGCGCGCAAGAAAGAGGCGCCAAAGCAACCCAAAGCCGAAAGAAACCA 7917
Qy 2721 AGAAGCAGCAAGCCAGAGAGCAAAAC---GCAAGCTTAAACAGAGGAAAGCAACGTA 2777
Db 7918 AAACACAGAGAAAGAAAGAAAGCACTCTGCAAAACCCAAACCCGAAAGAGACAACTGA 7977
Qy 2778 TGTGTATGAAGTTGAGTCCGACAAGACATTTCCGATC---ATGCTGAACGCGCAAGTGA 2834
Db 7978 TGGCACTCAAGTTGAGGCGCGACAGACTGTTCGACGTCAAAAATGAGGACGGAATGTCA 8037
Qy 2835 ATGATATGCTGTGCTGTGCGAGAGAAAGCTGATGAAACCTCAAGCTTGAAGAAAAA 2894
Db 8038 TCGGGACGCACTGGCCATGGAAGAAAGTAAATGAAACCACTCAAGTGAAGAACTA 8097
Qy 2895 TTGATATGAGCAATTAAGCGCGCTGAAATTGAAGAGGCTAGCATGACACTTTGAGT 2954
Db 8098 TTGACCAACCTGTGTCTCAAAAGCTCAAAATTCACAAAGTGTGACATAGACATGAGT 8157
Qy 2955 ACGGCGAGCTTCCCAAGAAATGAATCAGACGCTGCACTACCAAGCAACCAACAC 3014

Db 8158 TCGCACAGTTGCGGCTCAACATGAGAGAGGCGTTCACTTACACCAAGCAACCCCTG 8217
Qy 3015 CGGCGCTTCAACAATGGCACACGCGGCGAGTCCAGTATGAAATGGGAGATTAACCGTAC 3074
Db 8218 AAGGTTTAACTGACACACGAGACCGGTGCAATGATGAGATTTAAATTCATCC 8277
Qy 3075 CGAAGAGTGGGCGGGAAGGCGCACAGCGGAAAGACCGATCTTGACAAACAGAGCAGAG 3134
Db 8278 CCGCGGAGTAGGAGGAGAGAGAGACAGTGTGTGCTGCAATTAATGATTAATCAGGCCGG 8337
Qy 3135 TTGTGCTATTTGTTCTAGAGAGTGCAAAATGAGGACCGCTACGCGCTTTCACTGTCA 3194
Db 8338 TTGTGCGATAGTCTTCGAGAGGCGCTGATAGGGAACAAACCTGCCCTTTCGCTCA 8397
Qy 3195 CTGGAAACAGAAAGGGGTACCATTTAGGATACCCCGAAGGTTGAAACCGGTG--- 3251
Db 8398 CTGGAATGCAAAAGGGAAGACATCAAGCAACCCGGAAGGGAACAGAAAGATGTGTG 8457
Qy 3252 -----CACTAGTTACAGCGCTATGCTGTCTTGAATGTCAAGTTCCATGCAACAAAC 3305
Db 8458 CAGACACACTGTGTCAGCGCATGTGCTTGGAAACGTGAGCTTCCATGCAATGCGC 8517
Qy 3306 CACCGTGTCTATTCTACACGCGCAAGACAACTGACGTGCTGAAAGAAAGCTG 3365
Db 8518 CGCCACATGCTTACCCCGGAACATCCAGAGCTTTGACATCTTGAAGAGAAACGTGA 8577
Qy 3366 ACAATCCAAATTTGACAGCGCTGTGGAGAACGTTGAAT---GTCCATACAGCGCGC 3422
Db 8578 ACACAGAGCTTACAGAACCTGTCTCAAGCCCATTTGCGGTGCGATCTGCGGCGAA 8637
Qy 3423 CCAACAGAACATTAACGATGACTTCAACTGACAGTCCCTTACCTGGGCTTTCGCCGT 3482
Db 8638 GCAAAAGACGTCACTGACGACTTATCTTGAACAGCCGCTACTTGGGCAATGCTGT 8697
Qy 3483 ATTGCAACACTCAACCGCGTGTTCAGCCCAATAAAAATTGAGAAAGTGTGGAGCAAT 3542
Db 8698 ACTGTCAACATCTGAACCGTGTCTTGAAGCCGATTAAGATCGAGAGCTGTGGATAG 8757
Qy 3543 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3602
Db 8758 CGGACGCAACACCATGACATGACATGACATGACATGACATGACATGACATGACATGAC 8817
Qy 3603 CTGCGATGTCACCAATTCCTGTTACATGCTTTCGACCAAGCAACATGATCAAGAGAG 3662
Db 8818 CAGCAACTCAATTAAGTACCGCTACATGTGCTGAGAGAGATCAATACGTCAAAGAG 8877
Qy 3663 ACGATGAGAAATATGCTATGACGACATCTGGAACCTTCGCTGCTTGGCCAAAG 3722
Db 8878 GCACTATGATGATCATCAAGATGACGACCTCAGAGCCGTGTAGAAAGCTTAGCTAACAA 8937
Qy 3723 GGTACTCTGTTAGCTCAATGTCTCAGGTGACAGGTGAACGTCAGTATCACAGCG 3782
Db 8938 GATACCTTCTCTCGGCAAGTGTCTCAGAGGAGACGCTGACGTTAGTATGACGAGTA 8997
Qy 3783 GAGACTGAGAAATTCATGACCGGTGAGAAAAAGATCAGAGAGAAATTTGTCGTAGAG 3842
Db 8998 GCAACTGAGCAAGTATGACAAATGACCGGCAAGATTAACAAATTTGTGGAGCGG 9057
Qy 3843 AGGAGTACTTGTTCACACCGCTCAGTGAAGAGTGTGAAGTGCACGCTTTAGATCACT 3902
Db 9058 AAAAATATGACCTTACCTCCGTTACCGTGAAGAGATTCCTTTCACAGTGTACGACCTC 9117
Qy 3903 TGAAGAGAGCTGCGCGGGGTACATTAACATGACAGGCGCAGGCGCCACAGCGGTATAGT 3962
Db 9118 TGAAGAAACAAACCGCGCTACATCACTATGACAGGCGCGGAGCGCAAGCTTATAGT 9177
Qy 3963 CTTATCTGAGAGAGCTCAGCGGAAGTGAATTAAACCACTTCTGCGAAGACGTCA 4022
Db 9178 CTTATCTGAGAGATATACAGGAAAGTCTAACGGAAGCAACCATCGGAAAGAACTTA 9237
Qy 4023 CTTACGAATTAAGTGTGCGACTACAGCAGGTATCTGAGACCGGACGGAACGAAGTGA 4082
Db 9238 CGTACGAGTCAAGTGTGCGGATTAACAAAGACCGGTACCGTTAGCAGCCGTAACGAATCA 9297

QY	4003	ACGGCTGACCTAAAGCAAAAACAGTGCATTTGGCTTCAAGAGCCACCAACGAATGGGTCT	4142
Db	9298	CGGGCTGACCGCCATCAAGAGAGTGGCTGGCTTAAGACCGACCAACGAATGGGTCT	9357
QY	4143	TCAACTCGCCGGATCTTATTTAGGACACAGACCACTCAGTGCAGAGTAAATTCATATTC	4202
Db	9358	TCAATTCGCCGGACTTGATTCAGATATGCCACACACGGCCCAAGGAATTCATTTAC	9417
QY	4203	CATTCCGCTTGACACCGACAGTCTTGCCCGGTTCCGTTAGTCAACGCCCTTACGTACGA	4262
Db	9418	CTTTCAAGCTGATCCCGATGCTCCTGCATGGTCCCTGTGGCCACGGCCGAAGGTAGTAC	9477
QY	4263	AGTGGTTCAAGGGATCAACCCCTCCACCTGACCTGAAATGCGAACCAATTTGTCACACGA	4322
Db	9478	ACGGCTTTAAACATCAAGCTTCCATTAGACACAGCACCTGACATTTGCTCACACCA	9537
QY	4323	GAAAAATGGGGCTGCGAGCAGACCGCAACAGCAGATGATTAAGAGGTCTACATCCAGGA	4382
Db	9538	GGAAGCTAGGGGGCAAAATCCGGAAACCACTACTGAATGATTCATCGGAAGAAGCTTAGAA	9597
QY	4383	ATTTTTCTGTGGGGCGAAGAGGCTGGAGTACGTAATGGGTAACATGAACCAAGTCAGAG	4442
Db	9598	ACTTCACCGTGCAGCCGAGATGGCTGTGAATACAATATGGGCAATTCAGAACCCGTAAGGG	9657
QY	4443	TCTGGGGCCAGAGATCGGCACACAGCCACCACTGATGGCCGCAATGAGATCATATCC	4502
Db	9658	TCTATGCCCAAGATCTTGACACAGGAACCTCTCAGATGGCACACAGAAATAGTACAG	9717
QY	4503	ACTATATCATCGGCATCCAGTCTACACTGTCAATTTGCTGTGTGTGCTCTTGTGCTA	4562
Db	9718	ATTACTACATCGGCATCTGTGTACACCAATTTAGCCGTGATCAGCTGCTGTGGCA	9777
QY	4563	TCCGTGATGGACATCGCATCATCAGCAGCTTGCAATCGCCAAAGCAAGAAGATGCTGTA	4622
Db	9778	TGATGATTTGGCGTAACTGTGGACCAATTAATGTGCTTTAAAGGCGCCGTGATGCTGA	9837
QY	4623	CGCCATACGCGCTTGACCGAACCGCAACGGTACCACAGCATTAAGCGGTTTGTGCTGA	4682
Db	9838	CGCCATATGCTGCTGGCCCAAAATCCGTGATTCCACTTCGCTGGCACCTTTGTGCTGG	9897
QY	4683	TTTCGGCCACCAAGCTGTGAACATTTTGGAGAACTTTGAACATCTGTGTTTAACACCC	4742
Db	9898	TTAGTGTGGCTAAATGCTGAACATTTCAACGAGACATGAATTTCAATGTGTCAACAGCC	9957
QY	4743	AACGGTTCTCTGGGGCAGTTGTGCATTTCTCTGGCAGCGCTGTATTTCTTCCGCT	4802
Db	9958	AGCATTCCTCTGGGTCCAGCTGTGTATTCCTCGGCCGCTGTCAATGCTTCTAATGCGT	10017
QY	4803	GCTTTCACTGTCATGCTTTTATTTATGTGTGACGGCGTCTGCTGGGAGATGAGAGC	4862
Db	10018	GTTGCTCATGTGCTGCTGCTTTTATGTGTGTCGGCGGCTCTACCTGGCAGATGAGAGC	10077
QY	4863	CCTTCGAACATGCGACCACTGTGCAAAATGTTCCGGGGAATCCGGTAAAGGCGTGGTGG	4922
Db	10078	CCTACGAACATGCAACACTGTTCCAAATGTGCCACAGATACCGTAAAGCACTTGTGG	10137
QY	4923	AACGCGCAGGTTACGCGCACTTAACCTGAGATCAAGTGTCTCATCGAATTAACAC	4982
Db	10138	AAAGGGCAGGGTACGCCCCGCTCAATTTGGAGATTACTGTCAATGTCTCGGAGGTTTGG	10197
QY	4983	CTTCAACTAACAGAGATAGTGAACCTTGCAAAATTTCAACAGTCAATTTCTTACCAAG	5042
Db	10198	CTTCCACCAACCAAGAGATCAATCCCTGCAAAATTCACATGTGTGTCCCTCCCTAAAG	10257
QY	5043	TTAATATGCTGGGATCTCTCGAGTGCAAAGCATCTCAAGGGGGAATTAACATGTCGGG	5102
Db	10258	TCAATATGTGGGCTCTCTTGGAAATGTACGCCCGCGCTCACGAGACTAATCTGCAAG	10317
QY	5103	TTTTTGGCGGTGTACCTTTTCAATGTGGGAGGCGCAATCTTCTGTGACATGAGA	5162
Db	10318	TCCTTTGAGGGGTGTACCCCTTCAATGTGGGAGAGGACCAATCTTTTGTGACATGAGA	10377

QY	5163	ACACAAAC	AGTGAAG	CGCTAC	GTGAGT	TCCCTC	CAAC	CTGAC	TATATG	ATAC	GCAG	5222															
Db	10378	ACACC	CGAATG	AGTGA	GGCGTAC	GTGAA	TTGTCA	CGATG	CGGAC	TGAC	CAAGCCG	10437															
QY	5223	TCGC	ACTAAAA	AGTTCA	CA	CA	AGCTG	CTTG	AAAGT	CGGCG	CTATAG	ATAC	GCAG	5282													
Db	10448	AGCG	AGTTAA	GGTGC	ATATCT	GGCCG	CATG	AAATG	AGATAC	CTAG	CTATAT	AGTGT	ACGG	5282													
QY	5283	CCAC	CGCG	CACT	GTGAT	ACG	TTTTGT	CA	ATG	CGCT	CA	CGC	AGG	TTTCT	CA	CGG	CACT	GA	5342								
Db	10498	CTAC	CA	GTTC	CTT	CT	AGAT	GTG	AT	CGTGA	ACG	AG	CT	CA	CCAG	GA	CGT	CT	AA	GA	CT	GA	5342				
QY	5343	AGG	CTAT	AG	CA	GGG	CGCAT	AT	CA	AG	CGG	CTTTT	CA	CC	CTTT	TG	AC	CT	TA	AG	GT	GT	CA	TA	CA	5402	
Db	10558	AA	GTCA	TAG	CGT	GA	CCAA	TTTC	AG	CA	TCG	TTT	CA	CA	CA	TTTC	GA	TC	CA	AA	GG	TC	GT	AT	TC	10617	
QY	5403	GA	AA	GGG	CGT	GT	TTA	CA	CT	CA	TA	GA	CT	CC	CT	GA	AT	AG	AT	AG	AT	AG	AT	AG	AT	AG	5462
Db	10618	AT	CC	GGG	CGT	GT	TTA	CA	CT	CA	TA	GA	CT	CC	CG	GA	AT	AG	AT	AG	AT	AG	AT	AG	AT	AG	10677
QY	5463	TC	GG	CGAT	TAT	TC	AA	GA	CT	CT	CG	CT	GT	AT	GC	TA	CA	GA	CA	TAG	AT	AG	CC	CA	CT	GA	5522
Db	10678	TT	GA	GA	CA	TT	CA	AG	CT	CT	CG	CT	GT	AT	GC	TA	CA	GA	CA	TAG	AT	AG	CC	CA	CT	GA	10733
QY	5523	GG	CT	CT	GA	AA	CG	CTT	CT	GT	CA	GA	CA	CT	CC	CT	TA	CA	CC	CA	GA	GT	AT	CA	GG	GT	5582
Db	10738	GAC	TACT	CA	AG	CCCTT	CG	CA	AA	CA	CGT	GA	TC	CC	GT	TA	CA	CA	CG	AG	CG	GA	TC	GT	AG	AT	10797
QY	5583	AT	GA	AT	GT	GA	AA	CA	CT	CA	GG	AG	CA	CC	CT	GA	AA	GA	CA	GA	CA	AT	TT	GA	TC	TA	5642
Db	10798	TC	GA	AT	GT	GA	AA	CA	CT	CA	GG	AG	CA	CC	CT	GA	AA	GA	CA	GA	CA	AT	TT	GA	TC	TA	10857
QY	5643	AA	AT	TT	GA	AT	GT	GA	AG	CA	CC	CT	GA	AA	CA	GA	CA	AT	TT	GA	TC	TA	5702				
Db	10858	AG	AT	TT	GA	AT	GT	GA	AG	CA	CC	CT	GA	AA	CA	GA	CA	AT	TT	GA	TC	TA	10917				
QY	5703	TT	GA	CA	AT	CC	CT	GA	TG	CA	AG	TAT	GT	GA	T	CA	T	CA	T	CA	T	CA	5762				
Db	10918	TC	GA	CA	AT	CC	CT	GA	TG	CA	AG	TAT	GT	GA	T	CA	T	CA	T	CA	T	CA	10977				
QY	5763	GT	CC	CA	AG	T	GA	CA	CT	GA	AT	TT	GT	GA	T	CA	T	CA	T	CA	T	CA	5822				
Db	10978	AA	TT	GA	T	GT	GA	CA	CT	GA	AT	TT	GT	GA	T	CA	T	CA	T	CA	T	CA	11037				
QY	5823	ACA	AA	AG	CT	GA	CA	GG	AG	GA	CA	AT	TT	GT	GA	T	CA	T	CA	T	CA	5882					
Db	11038	AT	GT	AT	CC	GA	CC	GA	AG	GA	CA	AT	TT	GT	GA	T	CA	T	CA	T	CA	11097					
QY	5883	AG	GA	AG	CA	CA	CA	T	GT	GA	CT	CC	GT	GA	CA	T	CA	T	CA	T	CA	5942					
Db	11098	AA	AG	AT	CC	GA	CA	CT	GA	CT	CC	GT	GA	CA	T	CA	T	CA	T	CA	11157						
QY	5943	GCC																									

|||||
Db 11458 GCATTAATGATCAGGCTGTATATTAGTATCCCGCTTACCGCGGCAATATAGCAACACC 11517
|||
Qy 6269 AAAAAGTGAATGATTTCCGAGGAGACACAGTGCATTAATGCTGTGCAGTGT 6318
|||||
Db 11518 AAAAAGTGAATGATTTCCGAGGAGACACAGTGCATTAATGCTGTGCAGTGT 11567
|||||

RESULT 13

US-09-102-248-4
Sequence 4, Application US/09102248
Patent No. 6008035
GENERAL INFORMATION:
APPLICANT: Johnston, Robert E.
APPLICANT: Davis, Nancy L.
APPLICANT: Simpson, Dennis A.
TITLE OF INVENTION: System for the in Vivo Delivery and
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
STREET: 1211 East Morehead Street
CITY: Charlotte
STATE: No. 6008035th Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/102,248
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/801,263
FILING DATE: 19-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-147
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 11717 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-102-248-4

Query Match 36.8%; Score 2386.4; DB 3; Length 11717;
Best Local Similarity 64.5%; Pred. No. 0;
Matches 3745; Conservative 0; Mismatches 1991; Indels 74; Gaps 9;

Qy 580 ACTGACGGATGAAAGCGGAGCGGTATATTTTCATGAGAAACAGGCCAAGTCACTTC 639
|||||
Db 5761 ACTGACCTAAGCGGGGTAGTGGTACATATTTTCAGCGGACACAGGCCCTGGGCACTTGC 5820
|||||
Qy 640 AACGGAATCAGTACGTCAATGTAACTACAAAGAACTATATTTGGAGCCCTGCATG 699
|||||
Db 5821 AATATGAGTCCGTTCTCAGATCAGCTTACAGAACCGACCTTGGACGCAATTTCTGG 5880
|||||
Qy 700 AGAAGTATAGCCCGCGCTCGATCTCGAAAGAGAAATGTACAGAAAGAACTGC 759
|||||
Db 5881 AAGCAATCTAGCCCGCGCTCGATCTCGAAAGAGAAATGTACAGAAAGAACTGC 5940
|||||
Qy 760 AATTATGCGCTCTGAAAGAAATGAAAGCAGATTCATCAATCAAGAAAGTAAATATGA 819
|||||
Db 5941 AGATGATGCCACCGAAGCCAAAGCAGGTACCAAGTCTAGAAAGTAAATATCAGA 6000
|||||

Qy 820 AAGCAATTAAGCGGAGCGAATCTATTTCTGATTTGGGCAATATCTATCATCAGAAATGA 879
|||||
Db 6001 AAGCAATTAAGCGGAGCGAATCTATTTCTGATTTGGGCAATATCTATCATCAGAAATGA 6057
|||||
Qy 880 ATCTGTGAGTGTACAGAGTCAATATTCCTGACCAATCTACTGTCACCGTAATTA 939
|||||
Db 6058 ATCAGCGAATGTATTAAGATCACTACCGAAGCAATGATTCAGAGGTATCCGG 6117
|||||
Qy 940 ACAAGTTTACATCTGAGAGAGTGGCGTTAAACGTGCACTTAATTCAGAGAAAT 999
|||||
Db 6118 CGAATCTACTGACCCAAAGTTTGTCTGATCTGTTCACCAATCTACTGATGAAAT 6177
|||||
Qy 1000 ACCCTACAGTACAGTATTTATTAACAGATGAATTCAGATGCTTACAGATGTTG 1059
|||||
Db 6178 ACCGAGAGTACATCTTATCAATACAGACAGATGATGCTTACAGATGTTG 6237
|||||
Qy 1060 ACGGCGATGCTGTCTGATTAAGCAACCTTTTGTCCGCTAAACTGAGAACTACC 1119
|||||
Db 6238 ACGGAGAGTGTGCTGTGATATAGCACTTTTGTCCGCTAAAGTTACC 6297
|||||
Qy 1120 CAAGAAGCATGCTATTTGCAAGCAAGATTAAGATGAGCCGCTCCATGCTATACGA 1179
|||||
Db 6298 CGAAGAAGCATGCTATTTGCAAGCAAGATTAAGATGAGCCGCTCCATGAGATGAGA 6357
|||||
Qy 1180 ATACATTAAGAAATGATTTGCTGAGCTCTAAAGAAATGCAAGTACCAATTC 1239
|||||
Db 6358 ACAGCTTGCAAAAGCTGCTATTTGCTGAGCTCTAAAGAAATGCAAGTACCAATTC 6417
|||||
Qy 1240 GAGAAATACCTGCTTATGATTTGCGGCGCATTTATGTTGATGTTTCAGAAATACCAT 1299
|||||
Db 6418 GTGATTTGCAACATCTGATCTGAGCGCATTTCAAGCTTGAATGTTTCAGAAATATCAT 6477
|||||
Qy 1300 GCATGATGATGATCTGGGATACCTTTGCGATTAACCTATTTGCGTACAGAGAACG 1359
|||||
Db 6478 GTATGACGATGATTTGGAGAGTTTCCGAAAGCCATTTAGATCTACTGAGTTCC 6537
|||||
Qy 1360 TTACGCAATATGATGCAAAAGCTGAAAGGCGGAAAGCAGCAGATTTTGCATATCTC 1419
|||||
Db 6538 TTACGCAATATGATGCAAAAGCTGAAAGGCGGAAAGCAGCAGATTTTGCATATCTC 6597
|||||
Qy 1420 ATATCTAATAACCGTTGACAGAGATACCAATGATCAATTTGTCATGATCTTAAAGAG 1479
|||||
Db 6598 ATATCTAATAACCGTTGACAGAGATACCAATGATCAATTTGTCATGATCTTAAAGAG 6657
|||||
Qy 1480 ATGTCAAGTATCTCCGCGGACGAAACATACAGAGAGCGCTTAAGTCAAGTATTC 1539
|||||
Db 6658 ACCTGAAAGTTACACCTGGGACGAAACATACAGAGAGCGCTTAAGTCAAGTATTC 6717
|||||
Qy 1540 AGGCTGAGATCCCTGTTGCTACAGCTTACCTTTGCGGATCTCAATCGGAAATAGTCCGTA 1599
|||||
Db 6718 AAGCCGAGAACCCCTGGGACGCTTACCTTTGCGGATCTCAATCGGAAATAGTCCGTA 6777
|||||
Qy 1600 GACTGAATGCGGCTTCTGCAAAATATTCATCTCTTTGCAATGTCAGCGGAAGATT 1659
|||||
Db 6778 GACTGAATGCGGCTTCTGCAAAATATTCATCTCTTTGCAATGTCAGCGGAAGATT 6837
|||||
Qy 1660 TTGATGAGTATTTGCTGAACATTTTCCACAGCGGACCAACCAATGTTGAAACCGACATCG 1719
|||||
Db 6838 TTGATGAGTATTTGCTGAACATTTTCCACAGCGGACCAACCAATGTTGAAACCGACATCG 6897
|||||
Qy 1720 CGTGTGTTGATTAAGGAGGAGAGAGCGCTATTCCTTTGCGGCTGATGATCTTTGAGG 1779
|||||
Db 6898 CTTGTGTTGATTAAGGAGGAGAGAGCGCTATTCCTTTGCGGCTGATGATCTTTGAGG 6957
|||||
Qy 1780 ACTTAGTGTGACCAACCGCTTTGATTTGATAGAGGCGGCTTCGCAATATCAAT 1839
|||||
Db 6958 ACTTAGTGTGACCAACCGCTTTGATTTGATAGAGGCGGCTTCGCAATATCAAT 7017
|||||
Qy 1840 CTGTGACCTTACCTACAGAGAGGTTTAAATTTGTCATGATGAAATCCGATATGT 1899
|||||
Db 7018 CCACCCATCTGCGGAGGATCCGTTTCAAAATTTGCGGAGGATGATGAAATCCGATATGT 7077
|||||

QY 1900 TCTTAACGCTGTTTGTCAACACAGTATCATATGATTTGCTAGCAGAGTACTACGTG 1959
Db 7078 TCCTCAACCTCTTTTGTCAACACAGTATCATATGATTTGCTAGCAGAGTACTACGTG 7137
QY 1960 AACGGTTTAAACGCTGAGGCTGCGGCTCTATGCGGACGATTAACATAGTGCATGGTG 2019
Db 7138 AGCGGCTTAAACGCTGAGGCTGCGGCTCTATGCGGACGATTAACATAGTGCATGGTG 7197
QY 2020 TCGTCTCCGACACCTTGTATGCGGAGAGATGCGGACCTTGCTGGAACATGAAAGTAAAA 2079
Db 7198 TAGATCTGACAAAGAAATGCTGAGAGGTGTGCCACCTGCTCAACATGAGGTTAAGA 7257
QY 2080 TTATGATGCACTTATGCTATCAAGACCCCTACTTCTGTGGGGGATTTATCTGTGG 2139
Db 7258 TCATGAGCGAGTCTGCGCGAGAGACCGCTTACTTCTGCGGTGATTCATCTTGCAAG 7317
QY 2140 ACCAATTAACAGGACAGCTGCAAGATGTCAGACCTCTTAAAGGCTTTTAAAGCTTG 2199
Db 7318 ATTGGTTAATCTCCACAGGCTGTGCGGCGGACCCCTTGAAAGGCTGTTTAAAGTTGG 7377
QY 2200 GAAACCAATGGCGATGCTGATACCAAGACTGCGACCGCGCGGCGCATGCAATGATG 2259
Db 7378 GTAAACCGCTCTCCAGCCGACGAGCAAGACAAAGACAAAGACGCGCTGTGCTAGATG 7437
QY 2260 AAGCAATGCGATGAAACAGAAATTTGAAATTTACGAGATTTAGTGAAGGCGGTAGATCA 2319
Db 7438 AAACAAAGGCGGTGTTTAAAGTATTAACAGACACTTACAGTGGCGGTGGCAATC 7497
QY 2320 GATAGCAATCTACTGCGAGGCTGATCATGCTCTCTGTCCAGTTAGCCGAAAGCG 2379
Db 7498 GGTAGAGATGAGAACATCAACACTGCTGCTGCGCATTTGAAACCTTTTGGCCGAGACA 7557
QY 2380 TTAAGACCTTCAAGCATTAAGAGGAGGCCAATCACTCTGACGGCTGACCTTAATAGG 2439
Db 7558 AAAGGACATTTCAAGCCATCAAGGAGAAATTAACATCTCTACGGGTGCTTAATAGT 7617
QY 2440 TGACCTAGTA-----GACAGCACTTACCCACCGCGAGATGTTTCCATACC 2486
Db 7618 CAGCATACACATTTCACTGTGACTAATACCAACACCAACCACTGATAGAGATTCT 7677
QY 2487 CTCAGCTGAATTTTCCACAGTTTACCTTCAAAATCCGATGCTTACCGAGATCCAAAC 2546
Db 7678 TTAACATCTGCGCGCGCGCCCTTCCGCGCCCTCCACTGCGCATGTGAGGCGCGAGAA 7737
QY 2547 CTCCTAGGCGCGCTGAGGCGCGCTTTCGCGCCCGCTGCTCAATTCAGATCTTCA 2606
Db 7738 GGAGCAGGCGCGCCCGATGCTGCGCCGATGAGCTGCTTCCAAATTCAGCAATCTGA 7797
QY 2607 GGAGTCTGATGATCACTTGAATTTCAACA-----ACGATCACCTAATCCGCGCGAG 2660
Db 7798 CCACAGCGGTGAGTCCCTAGTCACTTGGACAGGCAACTAGACTCAAAACCCACGCCAC 7857
QY 2661 GTCCACCGCCAAAGAAAGAAAGTGTCTCTTACGCCAAACCTTACCTGAGCTTAAAGA 2720
Db 7858 GCCCGCCCGCGCGCAGAAAGGAGCGCGCCAAACCAACCGAAGCGAAGAAACAA 7917
QY 2721 AGAAGCAGCAAGCCAAAGAGACGAAC---GCAACCTTAAACGAGGAAAGCAAGCTA 2777
Db 7918 AAACACAGAGAAAGAAAGAAAGAACTGCAAAACCCGAAAGAGACAACTGA 7977
QY 2778 TGTGATAGATGAGTGTGACAGACATTTCCGATC---ATGCTGAAGCGCCCAAGTGA 2834
Db 7978 TGGCACTCAAGTTGAGGCGCAGACAGCTGTTCGAGTCAAAAATGAGAGCGAGATGTCA 8037
QY 2835 ATGATATGCTGTGCTTGTGAGGAGAGCTGATGAACCACTCCACGTTGAAGAAAA 2894
Db 8038 TCGGCGACGCACTGGCCATGGAAGAAAGTATGAACCACTCCACGTTGAAGAACTA 8097
QY 2895 TTGATTAATGAGCAATTAAGCGCGCGGTGAATTTGAAGAAAGCTAGCATTTAGAGT 2954
Db 8098 TTGACCACTGCTGTATCAAACTCAAAATTTCAAAAGTCTGACAGATTAAGCATGGAGT 8157
QY 2955 ACGGCGAGCTTCCCGAGAACTGAATGACACGCTGCAGTACACGAGCAAAACAC 3014

Db 8158 TCGCAGATGTCGGGTCAACATGAGAAAGTGAAGCGTTCACTCAACGAGCAACCCCTG 8217
QY 3015 CGGGCTTTCAACATCTGACACCAAGCGCGCATGTCAGATATGAGAAATGGAGATTTACCCTAC 3074
Db 8218 AAGGCTTTTAAACATGCGGACCAACGAGGCGGTGACATATAGTGAAGTGAATTTTACATCC 8277
QY 3075 CGAGAGAGTGGGCGGAAAGCGGACAGCGGAAAGCCGATCCTTGGAACAAGAGCAGAG 3134
Db 8278 CCCCGGAGTGAAGAGAGAGACAGTGTGCTGCTCAATTTATGATTAATCTCAGGCGCGG 8337
QY 3135 TTGTGCTAATTTGTTCTGAGAGGTGCAATGAGGCGACGCGATGACGCGCTTTCAGTGTCA 3194
Db 8338 TTGTGCGATAGTCTCTCGAGAGGCGGTGATGAGGAAACAAGATCCTCTTCGCTGTCA 8397
QY 3195 CTGGAACCAAGAAAGGCTGACATTTAGGATATCCCGAAAGTTCTGAACCGTGT--- 3251
Db 8398 CCGGAATGACAAAGGGAAGCAATCAAGACACCCCGAAAGGGAACAGAGTGTGTCTG 8457
QY 3252 -----CACTAGTTACAGCGGTATGCGGCTTTGGAATGTCACTTCCATGCGACAAAC 3305
Db 8458 CAGACCACTGTGTACGCGCATGTGCTTGTGGAACGTGACCTTCCATGCAATCGCC 8517
QY 3306 CACCGTGTCTATTCACTGACGCGCAAGACCACTGACGTGCTCGAAGAGAACGTG 3365
Db 8518 CGCCACATGTCTTACACCGCGAACCATTCAGAGCTTTGACATCTTGAAGAGACGTGA 8577
QY 3366 ACAATCAATTTAGACACGCTGTGGAACGTCTTGAAT---GTCAATACCGCGGC 3422
Db 8578 ACCAGAGCGCTACGACACCTGCTCAACGCGCATATGCGGTGCGATGTCTCCGCAAA 8637
QY 3423 CCAACGAAAGATTAACGATGACTTCACTGACGACGCGCTTACCTGGGGTTTGGCGCGT 3482
Db 8638 GCAAAAGAGCTCATGACATCTTTTACCTTGACCGCGCTTACCTGGGACATGCTGT 8697
QY 3483 ATTGACACTTCAACGCGGTGTTTCAAGCCCAATTAATTAAGAACGCTGTGGAGCAAT 3542
Db 8698 ACTGTCAACATATGAAACCGTGTGTTAGCCGATTAAGATGAGAGAGTCTGGAGTGAAG 8757
QY 3543 CTGATGATGATTCGATTTGAATTCAGAGTCTGGGCAATTTGGCTCAATAGCAGGCA 3602
Db 8758 CGGACGACAAACATTCGATACAGACTTCCGCGCAGATTTGATACCAACAAAGCGGAG 8817
QY 3603 CTGCGGATGTCACCAATTCGCTTACATGCTTTTGCACAGACCATGACATCAAGAG 3662
Db 8818 CAGCAAGCTCAATTAAGTACCGCTTACATGCTGTGACAGATCAATCACTCAAGAG 8877
QY 3663 ACAGATGAGAAATAGCTATTCAGACATCTGACACCTGCGCTGTGCTGGCCCAAG 3722
Db 8878 GCATATGAGATGACATCAAGATCAGCACTGAGACCGTGTGAAGGCTTAGCTACAAAG 8937
QY 3723 GGTACTTCTGTGACTCAATGTCTCTCAGGATGACAGTGAACCGTCAATCAAGAGCG 3782
Db 8938 GATACTTCTCTCGGAAAGTGTCTCCAGGGGACAGGTAAACGTTAGTATGAGAGTGA 8997
QY 3783 GAGCATCTGAAATTCATATGACCGGTGAGAAAGATCAGAGAAAGTTTGTGGTAGAG 3842
Db 8998 GCAACTAGCAACGTATGACCAATGTGCGCGCAAGATTAACCAAAATTCGTGGACCGG 9057
QY 3843 AGGAGTACTTGTCCACCCGTCATGAAAGTGTGAATGTGCAACGTTTACGATCACT 3902
Db 9058 AAAATATGACATCTCTCCGTTTACGCTTAAGAAATTCCTTGCACAGTGTACACGCTC 9117
QY 3903 TGAAGAGACGTCTGCGGAGTACATTAACATGACAGGCGACGCCACACCGCTATTAAGT 3962
Db 9118 TGAAGAAACAAACCGCGGTACTATCACTATGACAGGCGCGGACCGCACTATAGT 9177
QY 3963 CCTATCTGAGAGAGCGTCAAGGCGAAGTGTATTAACCACTTCTGGCAAGAACGCTCA 4022
Db 9178 CTTATCTGAGAGATTCATGAGGAAAGTCTAGCGAAGCCACATTCGGAAGAAACATTA 9237
QY 4023 CCTACGATGTATGTGTGGCACTTACAGCAGGTTATCTGTAGACCGCAAGCAAGATGA 4082

D 9238 CGTAGAGTGAAGTGGCGATTACAAAGCCGGTACGTTACGACCCGTACCGAAATCA 9297
Q 4083 ACGGCTGCATTAAGCAAAACATGCAATGGCTTCAAGAGGACCAACGAAATGGGTCT 4142
D 9298 CGGGCTCACCGGCATCAAGAGTGGCTGCTTAATGAAGCAACGAAAGGGGTCT 9357
Q 4143 TCAACTGCGGATCTTATTAGGACACAGACCACTGAGTCAAGTAAATTGACATTC 4202
D 9358 TCAATGGCGGACTTATGACATGCGACCAACAGCCCAAGGAAATTCATTAC 9417
Q 4203 CATTCGCTTGACCAACGACAGTCTGCCGGTTCGGTTAGCTCACACGCTTACAGTCA 4262
D 9418 CTTTCAAGCTGATCCGAGTACCTGATGGTCCCTGTGGCCCAAGCCCGAAGCTAGTAC 9477
Q 4263 AGTGGTTCAAGGATACCCCTGACCTGACCTGAATGCCAACAATGGTGAACA 4322
D 9478 ACGGCTTTAAACATACATAGCTTCAATTTAGACAGACCACTGACATGTCTACCA 9537
Q 4323 GAAATTTGGGCTGCGAGACAGCAAGCAATGATTTACAGGGTCTCATCCAG 4382
D 9538 GGAAGCTAAGGGCAATCCGGAACCACTGATGATGATGGAAGAGCGTTAGAA 9597
Q 4383 ATTTTCTGGGGCGAAGAGGGCTGAGTACGATGGGTTAACCATGAACCACTCAAG 4442
D 9598 ACTTCAACGTCGACCGAGATGGCTGGAAATATATGAGGCAATCAAGAACCGGTTAAGG 9657
Q 4443 TCTGGGCGCAGAGTCCGACACAGCCCAATGATGGCCGATGATCATATCC 4502
D 9658 TCTATGCCAAGATCTGACACAGAGACCTTCAAGATGGCCACAGAAATATATACAGC 9717
Q 4503 ACTATATCATCGGATCTGACATCTGCTTACATCTGCTGTGAGTGGCTCTTGCTA 4562
D 9718 ATTTACTACATCGGCATCTGTGTACATCTTACGCGTGGCATCACTGCTGTGGCGA 9777
Q 4563 TCTGTAGGACATGATCATACAGAGCTTGCATGCCAAGCAAGAAAGACTGCTGA 4622
D 9778 TGATGATGGCGTAACTGTTGACAGATTAATGCTGTAAGCGCGCTGAGTGGCTGA 9837
Q 4623 CGCCATAGCGCTTGCACCCGAACCGTAACCCACAGATTAAGCGGTTTTGGTGGCA 4682
D 9838 CGCCATATGCCCTGCGCCCAATGCCGATTTCCAACTTGGCTGGCGCT 9897
Q 4683 TTCCGCAACCAACGCTGAAACATTTGGAGAACTTTGAACCATCTGCTTTTAAAC 4742
D 9898 TTAGTGGGCTTAATGCTGAACATTTACCCGAGACATGATTAATGCTGAAACAGCC 9957
Q 4743 AACGCTTCTGTGGACAGTTGTGATCTCTGTGGAGCGCTTGTATTTCTGCTGCT 4802
D 9958 AGCCATTTCTGTGGGTCCAGCTGTGTAATCCCTGGCGCTGTGATCGTTCTAATGGCT 10017
Q 4803 GCTTTTATGCTGATGCTTTTATTTGTTGATGAGGGTGTGCTGGGAAAGGTATACG 4862
D 10018 GTTGTCTATGCTGCTGCTTTTATTTAGTTGCTGGCGGCTTACCTGGCAAGGTAGCG 10077
Q 4863 CTTTGAACATGACACCACTGTCGCAATTTGTCGGGGATCCGTAATGAAGCGTTGGCG 4922
D 10078 CTTAGCAATGACACCACTGTTCCAAATGTGCCAGATACCGTATAGGACATTTGTTG 10137
Q 4923 AACGCGAGTTAGCGCGCACTTAACCTGAGATCAAGGCTGTCTCATGGAATTAAC 4982
D 10138 AAAGGGGAGGGGTAGCCCGCTCAATTTGAGATTAATGTATGCTCCGAGGGTTTTGC 10197
Q 4983 CTTTCACTAACAAGAGTACGTAACCTGCAATTTCCACAGCTATTTCTTACACAG 5042
D 10198 CTTTCAACCAACCAAGATACATCTGCAAAATTCACACATGTGTCTCCCTCCCTAAG 10257
Q 5043 TTAATGCTGGGGCTCTGAGTGAAGGATCTTCAAGAGCGATTAACATGCGCG 5102
D 10258 TCAATGCTGGGCTCTTGGAAATGTCAAGCCCGCTGACGACAGATTAATCTGCAAG 10317
Q 5103 TTTTGGGAGTGTATCCCTTTCATGTGGGAGGCGCAATGCTTGTGTACAGTGA 5162
D 10318 TCTTTGAGGGGTGTATCCCTTTCATGTGGGAGGAGCAAAATGTTTTGGCAAGTGA 10377

Q 5163 ACAACAAGTGAAGGCGGTACGAGTTCCCTCAGACTGACATATGATCAGCAG 5222
D 10378 AAGCCAGATGATGAGGCGGTACGTGAATTTGACAGATTTCCGACCTGACACCGCG 10437
Q 5223 TCGCACTAAAAGTTCAACAGCTGCTGTGAAGTGGCGCTGCTATATGATACGCAACA 5282
D 10438 AGCGATTAAGTGTACATGCGCGATGAAGATGAAGTATATGTTAGTGAAGGAAACA 10497
Q 5283 CACCGGACCTGATACCTTTGTCAATGGCTTACCGCAGGTCTTCCACGGGACCTGA 5342
D 10498 CTACAGTTTCTAATGTTGTACGTGAAGAGTCAACAGGAAGCTTAAAGACCTGA 10557
Q 5343 AGGTATAGAGGAGCGATATACCGCTTTTTCACCTTTGACCAATGAAGCTGATCA 5402
D 10558 AAGTATAGCTGACCAATTTTACGATCGTTTACACATTTGATCAACAGTGTATCC 10617
Q 5403 GAAAGGGCTTGTATTAACAATACGACTTCCCTGATATGAGGCTATGAACCAAGAGCGT 5462
D 10618 ATGGCGGCTGTGTATCAACTATGATCTCCGGAATTCGAGGCGATGAACCAAGAGCGT 10677
Q 5463 TCGCGATATTCAAGATCTCTGCTGATGCTTACAGACATAGTACCCGACCTGACATAC 5522
D 10678 TTGGAGCATTTCAAGCTACCTCTTGACTAGCAAGATCTCATGCGCAGACAGACATTA 10737
Q 5523 GGGCTGAGAGCGCTTGTGTGAAGAACATCCAGCGCTCCCTACACCAAGCAGTATCAGGGT 5582
D 10738 GACTACTCAAGCTTTCCGCGCAAGACGTGATGCTCCGTACACAGAGCGCATCTGGAT 10797
Q 5583 ATGAATGTGAAAGAACATCTACAGACGACCCCTGCAAGAAAGACCAATTTGATGTA 5642
D 10798 TCGAGATGTGAAAGAACATCTACAGGCGCGCCATCTGAGAAACCGGCCCTTTGGGGTGA 10857
Q 5643 AATTTGAATGAGAGCTCTGCGAGCGCTTATCTGTGTTACGGGACATCCCTATCTGCA 5702
D 10858 AGATTTGAGTCAATCCGCTTTCAGCGGTGAGCTGTATAGGAAATTTCCATCTCTTA 10917
Q 5703 TTGACATCCCTGATGACGCTTTTGTGATCATCAGATCAACCAAACTTTTGAAGTTA 5762
D 10918 TCGACATCCGAAACGCTGCTTTATACAGACATCAATGACACACATGCTTCAACAGTCA 10977
Q 5763 GCTGCAAGTACAGACATGCTATTTATTTCTGACAGCTTTGGTGTCTCTAATCTACAGT 5822
D 10978 AATGTATGTCAAGTATGATCTTACAGCGACCTTCCGCGGATGCTACCTGAGT 11037
Q 5823 ACAAGCTGACAGGAGGACATTTGTCACTTCCACTCCACAGACAGCTGTTTGA 5882
D 11038 ATGTATCCAGCCGGAAGGACCAATGCCCTGTATGATTCGATTCGAGACAGCAACCTCC 11097
Q 5883 AGGAAGGACACACATGATGACGCTGAGCAGACATTAACATTTTATAGCATGCA 5942
D 11098 AAGATGCAAGTTCATGCTCTGGAAGAAAGAGCGGTGAAGATCACTTACAGACCGCGA 11157
Q 5943 GCCCACAAGCAAAATTTATATGTTTCTGCTATGCGGCAAGATCCACTGCAATGCTGAAT 6002
D 11158 GCCCAGAGGGAATTTATTTATGATCGCTGTGTGTAAGAGCAACATGCAATGCAAAAT 11217
Q 6003 GTAAACCAACCGCGGACCAATTAATGAGAACCACTAAAGTGCACCAAGAAATTTCCAG 6062
D 11218 GCAAAACCAACGCTGACATATGCTGAGCAACCCGCAAAATATGACCAAGAAATTTCCAG 11277
Q 6063 CGGCACTTCAAAAACATCTTTGAACTGGCTGTGACAGTGTGAGGAGCATCATCCC 6122
D 11278 CGGCACTTCAAAAACATCTTATGAGTGTGCTGACGCTTATGCTTATTAACACAGTAT 6182
Q 6123 TCAATTTATAGGACTTATATGATTTTGTGCTGACGCTTATGCTTATTAACACAGTAT 6182
D 11338 TATTAATTTATAGGACTTATATGATTTTGTGCTGACGATGATGCTGACACAGAAAT 11397
Q 6183 GACTGAGCGGAGACCTGA-CATAGCGTAAACTGATGATCTTCCGAGAAAGCTGTGT 6241
D 11398 GACCGCTACGCCCCCAATGAACCCGACAGCAAAACTCGATGATCTTCCGAGAAACTGATGT 11457

Qy 6242 GCATATCCACGCCGCT-----TGACACT 6268
 Db 11458 GCATATCCACGCCGCTGATATATAGATCCCGCTACCGCGGCAATATAGCAACC 11517
 Qy 6269 AAAATTCATGATTTTCCGAGAACACATGCTAATGCTGTGACGT 6318
 Db 11518 AAAATTCATGATTTTCCGAGAACACATGCTAATGCTGTGACGT 11567

RESULT 14

US-09-367-764-4
 Sequence 4, Application US/09367764
 Patent No. 6583121
 GENERAL INFORMATION:
 APPLICANT: Johnston, Robert E.
 APPLICANT: Davis, Nancy L.
 APPLICANT: Simpson, Dennis A.
 TITLE OF INVENTION: System for the In Vivo Delivery and
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
 STREET: 1211 East Morehead Street
 CITY: Charlotte
 STATE: No. 6583121th Carolina
 COUNTRY: USA
 ZIP: 28234
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/367,764
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/801,263
 FILING DATE: 19-FEB-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Sibley, Kenneth D.
 REGISTRATION NUMBER: 31,665
 REFERENCE/DOCKET NUMBER: 5470-147
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-420-2200
 TELEFAX: 919-881-3175
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 11717 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-09-367-764-4

Query Match 36.8%; Score 2386.4; DB 4; Length 11717;
 Best Local Similarity 64.5%; Pred. No. 0; Mismatches 1991; Indels 74; Gaps 9;
 Matches 3745; Conserved 0;

Qy 580 ACTGACGATGAGCGGAGCGTATTTCTCATCGGAAACAGCCAGGTCACCTTC 639
 Db 5761 ACTGACGATGAGCGGAGCGTATTTCTCATCGGAAACAGCCAGGTCACCTTC 5820
 Qy 640 AACGAAATCAGTACGTCAATGTAACAAAGAACTTATGATGGGCGCTCAGT 699
 Db 5821 AATGAGATCGCTTCTGAGATGAGTACGAAACCACTTGGAGGCGATGTTCTGG 5880
 Qy 700 AGAAGTATTAAGCGCGCGCTGATCTGAAAGAGAGAAATGTTACAGAAAGATGC 759
 Db 5881 AAGAGATTAAGCGCGCGCTGATCTGAAAGAGAGAAATGTTACAGAAAGATGC 5940
 Qy 760 AATATGCGCTCTGAGAAATAGAACAGGATCAATCAAGAAAGATGAAATATGA 819

Db 5941 AGATGATCCCAACCGAACCAACAAAGCAGGTACAGTCTAGAAAAGTAGAAAATCAGA 6000
 Qy 820 AAGCAATTAACGCGAGGAGCTCAATTTGGATTGGCAATATCTATCATCAGAGTGA 879
 Db 6001 AAGCAATTAACGCGAGGAGCTCAATTTGGATTGGCAATATCTATCATCAGAGTGA 6057
 Qy 880 ATTCGTGAGTGTATTAACAGTCAATTAATCTGTACCAATCTACTGCTCAACGTAATTA 939
 Db 6058 ATTCAGCCAGAAATGCTATAGATACACTACCCGAAACCAATGATATCCAGAGTGAACGG 6117
 Qy 940 ACAGGTTAATCTGCAAGAGTCCGGTTAAACGTCGAATTAATCAAGAGAT 999
 Db 6118 CGAATCTCTGACCCAAAGTTGCTGTAGCTGTGCAACAACTATCTGATGAGAAAT 6177
 Qy 1000 ACCCTAGTACGCAATGATATGTAAGATGAAATAGATGCGTATCTGACATGCTGG 1059
 Db 6178 ACCGACGATGAGCTTATATGATGATGATGATGATGATGATGATGATGATGATGATG 6237
 Qy 1060 ACGGCGATCGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1119
 Db 6238 ACGGCGATCGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 6297
 Qy 1120 CAAGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1179
 Db 6298 CGAAAGACACAGATATGAGCCCAACCTGCGATGCGGTTCCATCAGGATGACAGA 6357
 Qy 1180 ATACATTAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1239
 Db 6358 ACAGTTGCAAAAGCTGCTATGCTGCGGACTTAAAGAACTGCAACCTGCAACAAATGC 6417
 Qy 1240 GAGATTAATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1299
 Db 6418 GTGATTAATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6477
 Qy 1300 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1359
 Db 6478 GTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6537
 Qy 1360 TTACGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1419
 Db 6538 TTACGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6597
 Qy 1420 ATTAATCTAAACCGTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1479
 Db 6598 ATTAATCTAAACCGTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 6657
 Qy 1480 ATGTCAAAGTATCTCCGCGACGAAACATGACAGAGAGAGAGAGAGAGAGAGAGAGAG 1539
 Db 6658 ACGTGAAGTATCACTGCGACGAAACACACAGAAAGAAAGAGAGAGAGAGAGAGAGAGAG 6717
 Qy 1540 AAGCTGACGATCCCTTCTGACCGCTTACCTTTGCGGATCCATCGGAAATTAATGTCGTA 1599
 Db 6718 AAGCTGACGATCCCTTCTGACCGCTTACCTTTGCGGATCCATCGGAAATTAATGTCGTA 6777
 Qy 1600 GACTGAATGCGGATGCTTGTGCAATATCAATCTCTGCAATGTCAGGAGAGAT 1659
 Db 6778 GACTGAATGCGGATGCTTGTGCAATATCAATCTCTGCAATGTCAGGAGAGAT 6837
 Qy 6838 TTGATGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6897
 Db 1720 CGTGTGTTGATTAAGAGCAACAGCGATGCGCATTTGCGGCTGATGATGATGATGATGATG 1779
 Qy 6898 CCGTGTGATGATTAAGAGCAACAGCGATGCGCATTTGCGGCTGATGATGATGATGATGATG 6957
 Db 1780 ACTTAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1839
 Qy 6958 ACTTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7017
 Db 1840 CTGTGACATTAACGAGGAGAGAGATTAATTTGATGATGATGATGATGATGATGATGATGATG 1899
 Qy 7018 CCACCAATCTGCGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7077

QY 1900 TCTTAACGCTGTTTGTGCAACACACTAGTCAATATCATGATTGCTAGAGAGTACTAGCTG 1959
 DB 7078 TCTCAGCGCTCTTGTGTAACACAGTTCGATATGCTATATCGCAGAGAGTATTTGAGG 7137
 QY 1960 AACGGTTAAACCAACGTCAGCGCTGCGCGCTCTATTCGGCGAGCATTAATAGTGCATG 2019
 DB 7138 AGCGGCTTAAACGTCCTCAATATGTCAGCATTTATTCGGCGAGCATTAATAGTGCATG 7197
 QY 2020 TCGTCTCCGACACCTTGATGCGGAGAGATGCGCCACTTGGCTGTAACATGGAATGAAAA 2079
 DB 7198 TAGATATCTGACAAAGAAATGCTGAGAGTGTGCTCACTGGCTCAATATGAGAGTAAAG 7257
 QY 2080 TTATGATGAGATTATGATATGATCAAGACCCCTCTCTGTGGGGAGTTATCCTGATG 2139
 DB 7258 TCATATGCGAGTATGCGGCGAGAGACCGCTTACTTCTGCGTGAATCATCTTGAAG 7317
 QY 2140 ACCGATTAACAGGACAGCTGACAGATGCGAGACCTCTTAAAGCTTTTAAAGCTTG 2199
 DB 7318 ATTGCGTTACTCTCAACAGCGTGTGCGTGGGAGACCTTGAAAGGCTGTTTAAAGTTGG 7377
 QY 2200 GAAACCATTTGCGATGCGATGATATCCAGATCTGCGACCGCGCGGACCTGATATG 2259
 DB 7378 GTAAACCGCTCCCGACGAGCAGAGCAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 7437
 QY 2260 AAGCAATGCGATGGAACAGATTGGAATTAGAGAGATTAAGTGAAGAGAGAGAGAGAG 2319
 DB 7438 AAAACAAAGCGGTGTTTAAAGTATGATTAACAGACCTTAAAGAGTGGCTGTGCAATC 7497
 QY 2320 GATACGAGATCATCTGCGAGAGCTGATCATCATCTCTCTGCTACAGCTTAAAGAGAG 2379
 DB 7498 GGTATGAGTGAAGCAATCATCATCTCTCTGCTGATGAGAACTTTTGGCCAGAGCA 7557
 QY 2380 TTAAGAACTTCAAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2439
 DB 7558 AAAGAGCATTTCAAGCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7617
 QY 2440 TGAGGTGTA-----GACAGGACCTTACCCAGCGGAGAGAGAGAGAGAGAGAGAG 2486
 DB 7618 CAGATATGACATTTCACTGACTATATACCAACACCAACCACTGATATGAGAGATCT 7677
 QY 2487 CTCAGCTGAATTTCCACAGATTTACCTTACCAATTCGATGCTTACCGAGATCAAAAC 2546
 DB 7678 TTAAACATGCTGCGCGCGCGCGCGCTTCCGCGCGCCCACTGCGATGAGAGCGCGGAG 7737
 QY 2547 CTCCTAGCGCGCGCTGAGAGCGCTTGGCGCGCGCGCTGCTGCTCAATTCGAATCTTA 2606
 DB 7738 GAGAGCAGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7797
 QY 2607 GGAAGTGCATGATCACTTGAATTTCAAAAC-----ACGATACCTTAAATCCGCGCG 2660
 DB 7798 CCAAGCGCGTACGTGCTTACATTTGACAGGCACTGACCTTAAACCCCGACGCGCAC 7857
 QY 2661 GTCCACCG 2720
 DB 7858 GCG 7917
 QY 2721 AGAAGCAGCAAGCAAG 2777
 DB 7918 AAAACAG 7977
 QY 2778 TGTGATGAGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2834
 DB 7978 TGCGACCTCAATTTGAGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8037
 QY 2835 ATGATATGCTGCTGCTGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2894
 DB 8038 TCGGCGAG 8097
 QY 2895 TTGATTAAGAGCAATTAAG 2954
 DB 8098 TTGACACCGCTGTCTATCAAGCTCAAAATTCACCAAGTGTGTCAGCATATGAGAGT 8157

QY 2955 ACCGCGAGCTTCCCGAAGACATGAAATCAGACAGCTGACAGTACACGAGCGCAAAAC 3014
 DB 8158 TCGCAGAGTTCGCGGTCACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8217
 QY 3015 CGGCGCTTCAACATGCGACACAGCGCGAGCTCCAGATGAGAGAGAGAGAGAGAGAG 3074
 DB 8218 AAGGTTTCAACTGCGACACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8277
 QY 3075 CGAGAGAGTGGCGCGGAG 3134
 DB 8278 CCGCGAG 8337
 QY 3135 TTGCGCTATTTCTTACAGAGTGCATTAATGAGGCGAGAGAGAGAGAGAGAGAGAGAG 3194
 DB 8338 TTGTCGAGATGCTCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8397
 QY 3195 CTGGAACAGAAAG 3251
 DB 8398 CTGGAATAGCAAG 8457
 QY 3252 -----CACTAGTACAGCGCTATGCGTGTTCGAAATGTCAGTTCCTTGGAGACA 3305
 DB 8458 CAGCAGCAGTGTGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8517
 QY 3306 CACCGGTGCTATTTCACTGACCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3365
 DB 8518 CGCCGAGATGCTACAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8577
 QY 3366 ACAATCCAAATTAACAGACAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3422
 DB 8578 ACCAGAGCGCTACAGACAGCTGCTCAAGCCATATGCGGTGCGGTGCGCGAGAGAG 8637
 QY 3423 CCAAG 3482
 DB 8638 GCAAAAG 8697
 QY 3483 ATTGAG 3542
 DB 8698 ACTGTACAGATCTGAG 8757
 QY 3543 CTGATGATGAGATGATTAAGATCAAGTCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 3602
 DB 8758 CGAGCAGACACACATTCGATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8817
 QY 3603 CTGCGAGATGACCAATTCGATTAAGATCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAG 3662
 DB 8818 CAGCAG 8877
 QY 3663 ACAGATGAG 3722
 DB 8878 GCACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8937
 QY 3723 GGTATCTTCTGTTAGCTCAATGCTCTCAGGTGACAGTGTACGATACAGAGCG 3782
 DB 8938 GATATCTTCTCTGCGAG 8997
 QY 3783 GAGCATGAGATTAATGACAGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3842
 DB 8998 GCAACTGAG 9057
 QY 3843 AGAGATGCTGTTCCACAGCGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3902
 DB 9058 AAAAATATGACCTTACCTGCGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9117
 QY 3903 TGAAG 3962
 DB 9118 TGAAG 9177
 QY 3963 CCTATCTGAG 4022
 DB 9178 CCTATCTGAG 9237
 QY 4023 CCTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4082

Db 9238 CGTAGAGTGGCAAGTGGCCGATTAACAAGACCGGTACCGTTACGACCCGTACCGAAATCA 9297
Qy 4083 ACGGCTGACTAAACAAACACATGCTATGCTTAACAAGGCAACCAACGAAATGGGTCT 4142
Db 9298 CGGGTGAACCGCCATCAAGCAGTGTGCTCTAATGAAGGCAACCAAGAGTGGGTCT 9357
Qy 4143 TCAACTCCCGGATCTTAATTAAGGCAACAGACCACTCAGTGGCAAGTAATTAATGCAATTC 4202
Db 9358 TCAATTCGCCGACTTGAATCAGACATGCCGACCAACGCGCCCAAGGAAATTTGCAATTAAC 9417
Qy 4203 CATTCGCTTGAACACGACAGTGTGCGCGTTCCGTTAGCTTCAACGCTCAAGTACGA 4262
Db 9418 CTTTCAACCTGATCCCGAGTACTGATAGTGTCTCTTTGCCCCACCGCCGAAAGTATGAC 9477
Qy 4263 AGTGGTTAAAGGATCAACCTCCACCTGACCTGCAATGCAACCAATTTGCTGAACAAGA 4322
Db 9478 ACGGCTTTAAACATCAAGCTTCAATTAAGACACAGACCACTGACATTTGCTCAACACA 9537
Qy 4323 GAAATTTGGGCTGCGACAGACAGACAGACAGATGATTAAGAGGTCTACATCCAGGA 4382
Db 9538 GGAAGCTAGGGGCAATCCGGAACCACTACATGATGATCATCGAAAAGCGTTAGAA 9597
Qy 4383 ATTTTCTGTGGGCGAAGAGGCTGAGTACGATGGGGTAAACCATGAACCAATCAGAG 4442
Db 9598 ACTTCAACGTCGACCGAGATGGCTGGAATACATATGGGGCAATCACGAAACCGTAAAGG 9657
Qy 4443 TCTGGGCCCCAGAGTCCGACCAAGCCGACCAATGATGATGCGGATGATCATCATTC 4502
Db 9658 TCTATGCCCAAGATCTGACCAAGGAGACCTCAAGATGGCCACCAACAAATAGTACAGC 9717
Qy 4503 ACTATTTATCATGCGATCAAGTCTACATCTGATGTGTGTGTGTGTGTGTGTGTGTGTGTA 4562
Db 9718 ATTTACTACATATGCAATCTGTGTACACATTTAGCCGTGCAATCAGCTGTGTGTGTGCA 9777
Qy 4563 TCTGTGAGGCACTGCAATCATACGAGCTTGCAATGCCAAGCAAGAAAGACATGCTGTA 4622
Db 9778 TGATGATGGCGTAACTGTGTCGACATTAATGTCCTGTAAGCGGCGGTGATGCTGTA 9837
Qy 4623 CGCCATAGCCCTTGACACGAAACGAAAGGTAACCAACAGATTAAGCGTTGTGTGTGCA 4682
Db 9838 CGCCATATGCCCTGCGCCCAATGCGATTTCAACTTCGCTGCACTTTTGTGTGTGTG 9897
Qy 4683 TTCCGCAACCAACGCTGAAACATTTGGAAGAACTTTGAACCATGTTGTTTAAACAAC 4742
Db 9898 TTAGTCCGCTAATGCTAATTAACATTTACCGAGACCATAGATTAATTTGTTGCAACGCC 9957
Qy 4743 AACGTTTCTGTGGACAGTGTGCAATTCCTGTGAGCGGCTTGTATTTGTTCCGCT 4802
Db 9958 AGCCATTTCTGTGGGTCCAGGTGTATACCTGTGCGGTGTATCGTTCTAATGGCCT 10017
Qy 4803 GCTTTTATGCTGATGCTTTTATTTATGTTTGAAGGCGTGTGCTGTGGAAAGTGAAGC 4862
Db 10018 GTTGTATGCTGCTGCTTTTATTTATGTTTGAAGGCGGCTTACCTGTGCAAGTGAAGC 10077
Qy 4863 CTTTGAACATGCGACCACTGTGCAAAATGTTCCGGGATCCCGATTAAGGCGTTGTGCG 4922
Db 10078 CTTAGAACATGCGACCACTGTGCAAAATGTTCCAGATTAACGATTAAGGCACTTTGTTG 10137
Qy 4923 AACGCGAGTTTACGCGCACTTTAACTGAGATCAGCGTGTCTCATCGAATTTAACAC 4982
Db 10138 AAAAGGAGGGGTAGCGCCGCTCAATTTGAGATTAATGCTATGCTCTCGAGGTTTTCG 10197
Qy 4983 CTTCAACTTAACAGAGTACGATGCTGCAATTTCAACAGATCATTTCTTCAACCAAG 5042
Db 10198 CTTTCAACCAACCAAGATTAACATCTGCAATTTCAACCACTGTGTGTCTCCCTTAAG 10257
Qy 5043 TTAATGCTGGGGGTCCCTGAGTGAAGGATCTCTAAAGCGGATTAACATGATGCGCG 5102
Db 10258 TCAATGCTGGGCTCTTGTGAATGCAAGCCGCGCTCAAGCAGACATTAACCTGCAAG 10317
Qy 5103 TTTTGGGCTGTATCCCTTTCAATGTGGGAGGCGCAATGCTTGTGTGACATGAGA 5162

Db 10318 TCTTTGAGGGGTGTATCCCTTCATATGTGGAGAGGACCAATGTTTTGCGACATGAGA 10377
Qy 5163 ACACAACTAGATAGAGGCGGTAAGTGTGCTTCCGACATCTGCACTTAATCAACGAG 5222
Db 10378 ACAGCAATATAGTGAAGCGTACGTTGATTTGTACGAGATTTGCGGACTGACACGCGC 10437
Qy 5223 TCGCACTAAAGTTTCAACAGCTGCTGAAAGTGGCGTGTATGATATACGCAACA 5282
Db 10438 AGGCAATTAAGGTATCTGCTCCGATGAAAGTGAAGCTATGATGTATGATGAGGAACA 10497
Qy 5283 CCACGCGCACCTGATACGTTTGTCAATGAGCGTCAACCGAGTTCTTACGCGACTGA 5342
Db 10498 CTACAGATTTCTGATGTGTACGTAACGAGATCAACAGAACTGTCTTAAGACTGTA 10557
Qy 5343 AGCTATGACAGGCGGATATACGCGGTTTTTACCTTTGACATTAAGTGTGATCA 5402
Db 10558 AAGTATAGCTGACCAATTTGATGATGTTTACCACTTGTATCAAGATGCTGATATC 10617
Qy 5403 GAAAGGGGCTGTTTACACTACGACTTCCCTGAGTATGAGCTTGAACCAAGAGCGT 5462
Db 10618 ATCCGCGCTGTGTACACTATGATCTTCCGGAATACGAGCGATGAACCAAGAGCGT 10677
Qy 5463 TCGGATATTTCAAGATCTCGCTTGTATGCTACAGACATGATGACCGGACATGATAC 5522
Db 10678 TTGAGACATTTCAAGCTACTCTTGAACCAAGATCTCATGCGCAGACACAGATTA 10737
Qy 5523 GCTGTGTAAGCTTTTGTGTCAAGAACTCCAGTTCCTTACACCCAGACATTAAGGT 5582
Db 10738 GACTACTCAAGCTTTCCGCAAGAACGTGATGTCCCTACACGAGGCGCATCTGTAT 10797
Qy 5583 ATGAAATGTGAAGAACTACAGACGACCCCTGCAAGAAACAGACACTTTGATGTA 5642
Db 10798 TCGATATGTGAAGAACTACAGGCGCCCACTGACAGAAACGCGCTTTGCGGTGCA 10857
Qy 5643 AAATTTGAAGTGAAGCTTGTGAGCGCTTAACTGTCTTACGCGGACATCTCTGTA 5702
Db 10858 AGATTTGAGTGAATCCGCTTGTGAGGCGGTGAGCTGTCTTACGCGGAACATTTCCATCTGTA 10917
Qy 5703 TTGATATCCCTGATGAGCTTTTGTGATATCATGATATCAACCAATTTTGAAGTTA 5762
Db 10918 TCGACATCCCAAGCGCTCTTATACAGACATAGATGACACACTGTGTCTCAACAGTCA 10977
Qy 5763 GCTGCAAGTGAAGACTATTTATCTGAGACTTTGTGAGTTCTTAACTATGATAGT 5822
Db 10978 AATGTATGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11037
Qy 5823 ACAAGCTGACAGGAGGACATTTGTCCAGTTCACTCCACTCCACGACAGCTGTTTGA 5882
Db 11038 ATGTATCCGACCGGAGAGGACATGCTGTATCAATTTGCAATTTGACACAGAACCTTCC 11097
Qy 5883 AGGAAGGACCAACATGTGATGCTGCGGTAGGACAGATTAACCTATTTTGAACATGTA 5942
Db 11098 AAGAGTGAACGATTCATGCTGTGAGAAAGAGGCGGTGACATCTTACAGCACCGGA 11157
Qy 5943 GCCCAAGCAAAATTTATGATTTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 6002
Db 11158 GCCCAAGGCAAACTTTATTTATGCTGTGTGTGTAAGAAACAATGATGATGATGAT 11217
Qy 6003 GTAAACCAACCGCGGACCACTAATTTGAGAACCACTAAGTGCAGCCCAAGATTTCCAG 6062
Db 11218 GCAAAACCAACGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11277
Qy 6063 CGGCAATTTCAAAACATTTGGAATGCTGCTTTGCACTGTTTGGGGAGCATCATCC 6122
Db 11278 CCGGCAATTTCAAAACATTTGGAATGCTGCTTTGCACTGTTTGGGGAGCATCATCC 11337
Qy 6123 TCATTTGTTGAGCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6182
Db 11338 TATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11397
Qy 6183 GACTGAGCGGACACTGA - CATAGCGTAAACTGATGATGATGATGATGATGATGATGAT 6241
Db 11398 GACCGTATACGCCCAATGACCGACAGCAAAACTGATGATGATGATGATGATGATGAT 11457

Db 7003 CCACCATCTACCTACGGGTACTCGTTTAAATTCGGGGCGATGATGAATCCGAATGT 7062
Qy 1900 TCTTAACGCTGTTTGTCAACACACTAGTCAATATCATGATGCTGACAGACTACG 1959
Db 7063 TCCTCAACTTTTGTCAACACAGTTTGAATGTGTTATCGCCAGACAGATCAAGANG 7122
Qy 1960 AACGGTTAACACGTCAGCGTGGCGGCTCTATCGGCGACATTAACATAGTCAGTGT 2019
Db 7123 AGCGGCTTAAACGTCAGATGTCAGCGTTCATTGGCGACGACAAATCATACATGAG 7182
Qy 2020 TCGTCTCGGACACTTGTATGGGAGAGATGGCCACTTGGCTGAACATGAGTAA 2079
Db 7183 TAGTATCTGACAAAGAAATGCTGAGAGTGGCCACCTGGCTCAACATGGAGTTA 7242
Qy 2080 TTATGATGACGATTATGGTATCAAGACCCCTACTCTGTTGGGGATTTATCTG 2139
Db 7243 TCATGAGCGCATCATCGGTGAGAGCACCTTACTTCTGCGCGGATTTATCTTGCA 7302
Qy 2140 ACCAGATTAACAGGCAAGCCTGACAGTGCAGACCTCTTAAAGGCTTTTAAGCT 2199
Db 7303 ATTGGTACTTCCACAGCGTGGCGGTCGGATCCCTGAAGAGCTTTAAGTTGG 7362
Qy 2200 GAAACCATTTGCGACTCGATGATCCCAAGCTGCGACCGCGCGGCACTGATG 2259
Db 7363 GTAAACCCCTCCACCGACGACGACGAAGACGAAGACGAGCGCTCTGATGAT 7422
Qy 2260 AAGCAATGCAATGAAAGAAATTTGAATTAACGAGATTAAGAGGCGGATGATCCA 2319
Db 7423 AAACAAAGGCGGTGTTTGAAGATTAATTAACAGGCACTTTAGCAGTGGCGGAGCA 7482
Qy 2320 GATACGATCATCTAGCGAGGCTGATCATCAGCTCTCTGCAAGTTAGCCGAAG 2379
Db 7483 GGTATGAGTGAACAATTAATTAACCTGTCTACTGGCATTTGGCCAGAGCA 7542
Qy 2380 TTAAAGACTTCAAGAGCATTAAGAGGAGCCCAATCACTCTTACGCGCTGA 2439
Db 7543 AAAGGCAATTCACATCAAGAGGGAATTAAGCATCTTACGCTGCTCTTAAT 7602
Qy 2440 TGAGCTATGACAGGACCTAC-----CAACGGGAGAAATGTTCCATACC 2486
Db 7603 CAGCATAGTACCTTCACTGCTAATTAACACACACACATGATAGAGATCT 7662
Qy 2487 CTCAGCTAATCTTCCACAGATTTCACCTACAAATCCGATGGCTTACAGGATCCA 2546
Db 7663 TTAACTGCTGGCGCGCGCCCTTCCCGGCCCACTGCCATGTGAGAGCGCGGAA 7722
Qy 2547 CTCCTAGCGCGCTGAGAGCGGTTTCGCGCCCGCTGCTGCTCAATCGAAGCT 2606
Db 7723 GAGAGCAGGCGGCGCCGATGCTGCGCCGACGAGGCTGCTTCTCAAAATCCAG 7782
Qy 2607 GAGAGTCAATGATCACTTTGACTTTCAACACATCACTATCCGCGGCAAGTCCA 2665
Db 7783 CCACAGCGCTGACGCTCCTTATGATGACAGGCAATGACCTCAACCCCACTCC 7842
Qy 2666 -----CCGCAAGAAAGAGAGTGTCTTAAGCAAAACCTACTGACCTTAA 2720
Db 7843 GCCCGCCACCGGCGAGAGAGAGGCGCCCAAGCAACACAGAACCGAAGAAACCA 7902
Qy 2721 AGAAGCAGCAAGCCAAAGAGAGAAAC--GCAGCCTTAACAGAGAAACAGACGTA 2777
Db 7903 AAACGAGAGAGAAAGAAAGAAAGCACTGCAAAACCAAAACCGGAAAGAGACGCG 7962
Qy 2778 TGTGTATGAGTGTGAGTGCAGCAAGACATTTCCATCATG--CTGAACGCGCAAGTGA 2834
Db 7963 TGGCACTTAAGTTGAGCGCGACAGATTTGTCAGCTCAAGAACAGAGACGAGATGCA 8022
Qy 2835 ATGATATGCTGCTGCTGCGAGAGAGGCTGATGAACCACTTCCAGCTGAAGAA 2894
Db 8023 TCGGCGAGCACTGCGCATGAGAGAGATTAATGAACCTTGCACGTGAAGAAACCA 8082
Qy 2895 TTGATATGAGCAATTAGCGCGCGTGAATTGAAGAGGCTAGCATGATGCACTTGA 2954

Db 8083 TCGACACCCGTGTCTATCAAACTCAATTTACCAAGTGTGACATATCGACATGAG 8142
Qy 2955 ACGGCGAGTTCGCCAGAACATGAATCAGACACGCTGAGATACACAGCAAAACCA 3014
Db 8143 TCGCAGATTTGCGAGTCAACATGAGAGAGGCACTTCACTTACACAGATGAACCC 8202
Qy 3015 CGGGCTTTACACCTGGACCAACGCGGAGTCCAGTATGAGAAATGGAGATTTACG 3074
Db 8203 AAGATTTATTAATGCGACACAGGAGCGGTGACATATGAGAGATGATTTACATCC 8262
Qy 3075 CGAAGAGAGTGGCGGAGAAAGCCGACACCGGAACCGATCTTGAACAAGAGCAG 3134
Db 8263 CTGCGGAGTGAAGAGCAGAGAGACAGCGGTGCTCCATCATGATTAATCCGCTGG 8322
Qy 3135 TTGTGCTATTGTTTAGAGGTGCAATGAGGAGCGATGACGCGCTTTCAGTGTCA 3194
Db 8323 TTGTGCTATTGTTTCGTTGCGTGTATGAGAGAACAGATCCCTTTCGCTGCA 8382
Qy 3195 CTGGAACCAAGAAAGGAGTGAACATTAAGGATATCCCGGAAGTTCGAAACCGTGT 3251
Db 8383 CCGTGAATATGTAAGAGAGAACATTAAGAGACCCCGGAAGGACAGAGAGTGTCCG 8442
Qy 3252 -----CACTAGTTACAGCGCTATGCGTCTTGAATGTCACTTCCCATGCA 3305
Db 8443 CAGCACACTGCTACGCGCAATGTGTTGCTCGAAATGTGAGCTTCCCATGCGAC 8502
Qy 3306 CACCGTGTGCTATTCATGACGCGCAACGACGACGATGCTGCAAGAGAGCTGCG 3365
Db 8503 CGCCCAATGCTATACCGCGCAACCTTCAGAGCCCTGACATCTTGAAGAAAGTGA 8562
Qy 3366 ACAATCCAAATTTACACACGCTGCTGAGAAAGTGTGAATGCTATCAC--GCCGC 3422
Db 8563 AOCATGAGGCTAGATATCCTGCTCAATGCTATTCGCGGATGCTGTGCAAGAA 8622
Qy 3423 CCAAAGCAATTAACGATATTCACATGACGATCCCTTCTGCGGATTTGCGCT 3482
Db 8623 GCAAAAGAGGCTATGACCATTTTACCTGACGACGCTCTTCTGCGGACATCTG 8682
Qy 3483 ATTGAGACACTCAACGCGGTGTTACGCCCAATTAATGAGAACGTTGGAGCGAT 3542
Db 8683 ACTGCACATATGCTACGCTGCTTACGCTTCAATGAGAGAGGCTGCGAGAG 8742
Qy 3543 CTGATGATGATGCTATTAATGATTCAGAGTCTGCGCAATTCGCTTACATCAGGAGCA 3602
Db 8743 CGAGAGATTAACCATAGCATACAGACTTCCGCCAATTTGATGACACAAACGAG 8802
Qy 3603 CTGCGGATGTCACCAATTCGTTTACATGCTTTTTCGACACGACCATGATCAAGAG 3662
Db 8803 CAGCAAGCGCAAAACAAATACGCTTACATGCTTGAACAGATGACACCGTTAAAG 8862
Qy 3663 ACAATGAGAGAAATGATATGACATCTGAGACCTGCGCTGCTTGGCCCAAG 3722
Db 8863 GACCATGATGACATCAAGATTTAGCACCTGAGACCTGTAAGAGGCTTGTGCTA 8922
Qy 3723 GGTACTTCTGTTAGCTCAATGCTCTCCAGGATGACAGTGAACGCTGATACAGAG 3782
Db 8923 GATACTTCTCTCGCAAAATGCTCTCCAGGAGACGCTTAAAGCTTATGATGATGA 8982
Qy 3783 GAGCATCTGAAATTCATGACACGCTGAGAAAGAAAGATCAGAGAGAAATTTGCT 3842
Db 8983 GCAACTGACGAACGTCATGATGACATGCGCCGACAGATTAACAAATTCGTTGAG 9042
Qy 3843 AGAGTATCTTGTCCACCGCTCCATGGAAGCTGTGAAGTGCACGTTTACATCACT 3902
Db 9043 AAAATATGATCTACCTCCGTTACAGGTAATAAAATTCCTTGACACATGATGAC 9102
Qy 3903 TGAAGAGAGCTTCCGCGGTATCAATTAACATGACAGGCGCAGCGCTTAAAGT 3962
Db 9103 TGAAGAAACAACTGCGAGCTTACATCACTATGACAGGCGCAGACCGCTTATCAT 9162
Qy 3963 CCTATCTGAGAGAGCGTCAAGGAGATGATCAATTAACACCTTTGCGAAGAGCTCA 4022
Db 9163 CTAACCTGAAAGATCATCAGGAAAGTTTACGCAAAAGCCGCACTTGGAAAGCATTA 9222

OY	4423	CTTACGAAATGAATGTGGGCACTACAGCA	CAGTATCCGACACGACGAAAGAAATA	4082
Db	9223	CGTATGATGCAATGTGGGCACTACAGCA	CCGAAACCGTTTCGACCCGACCGAAATTA	9282
OY	4083	ACGGCTGCACCTAAAGCAAAA	CAGTGCAATTCCTTACAGAGCGACCAAGAAATGGGTCT	4142
Db	9283	CTGGTTGCAACGGCATCAAGCAAGTGCGT	CGTTTAAAGCGACCAAGAAAGTGGGTCT	9342
OY	4143	TTCAACTGCGGGATCTTATTAGGCACAC	AGCACTCAGTGCAGAGTAATTGCACATTC	4202
Db	9343	TTCAACTCACCGGACCTTATCAGACATGA	CCACACACACCGCCCAAGGAAATTCGCAATTCG	9402
OY	4203	CATTCGCGCTTGACACCGACAGTCTG	CCCGGTTCGGTTAGTCTCACGCGCTACAGTACGA	4262
Db	9403	CTTTCAAGTTGATCCCGAGTACTGCA	TGATGGTCCCTTGTCGCCAGCGCGGAATGTAATAC	9462
OY	4263	AGTGTTCAAAAGGCATCACCTCCACCTGA	CTGCAATTGCGACCAATTCGTCACACGA	4322
Db	9463	ATGGCTTTAAACATCATCAGCTCCCAAT	TATGATACAGACCACTTGACATTGCTTCACCA	9522
OY	4323	GAAATTTGGGGCTGCCAGACAGCCAA	CAGCAATAGATTTACAGGGTCTACATCCAGA	4382
Db	9523	GGAACTAGGGGGCAAAACCCGGAACCA	CACTAAATGATTCGTGGAAAGACGCTCAGAA	9582
OY	4383	ATTTTCTGTGGGGCGAGAAAGGGCTG	AGATGATCTATGAGGGAATACATGAAACAGTCAAG	4442
Db	9583	ACTTCACCGTGCACCGAGATGGCTGG	AAATACATATGAGGAAATCATGAGCCAGTGAAGG	9642
OY	4443	TTGTGGGCCCAAGAGTGGGCAACAGG	CGCACTCAATGATGCGCATCATATCC	4502
Db	9643	TCCTATGCCCAAGATCGACACAGGAA	ACCTCAGATGGCCACAGAAATGTAACAC	9702
OY	4503	ACTATTTACATCGGCATCCAGTCTACA	CTGTCATTTGCGTGTGGTGTCCGCTCTTGCTA	4562
Db	9703	ATTACTACATCGGCATCTGTGTACAC	CACTTTAGCCGTGCATACGCTACCGTGGCA	9762
OY	4563	TCCTGTGAGGCACTGATCATCAGCA	CTTGACCTTGACCCCAAGCAAGAAAGACTGCTGA	4622
Db	9763	TGATGATTTGGGCTTAACTGTTGAC	AGTGTATTATGTCCCTGTAAAGCGCCCTGATGCTGA	9822
OY	4623	CGGCATACGGGCTTGACCCGACGCA	ACGGTATACCAAGATTAGGGTTTGTGCTGA	4682
Db	9823	CGGCATACGGCCCTGTGGCCCAAG	CGCGGTAAATCCCAATTCGCTGGCACTCTGTGTGCTGG	9882
OY	4683	TTCCGCCAACCAACGCTGTAACA	ATTGTGAGAAACCTTTGAACCATCTGTGGTTTAACACC	4742
Db	9883	TTTAGTGGGCCAATGCTGAAACGTT	CAACGATGATTACTTGTGGTGAACAGTC	9942
OY	4743	AACGTTTCTGTGGGCAAGTTGTG	CAATTCCTCGGACGCGCTGTATTCTGTCCGT	4802
Db	9943	AGCGTTCTCTGTGGGTCAAGTTGTG	CAATTCCTTTGCGCTTTCATCGTTCAATTCGCT	10002
OY	4803	GCTTTATATGTGTGATACCTTTTT	TATTTGATGACGCGCTCTGCGGGGAAGGTGAGC	4862
Db	10003	GCTGCTCTGTGCTGCTCTCTTTT	TATTTATGTGTGCGGCGCTTACCTGTGGAGATGAGC	10062
OY	4863	CTTTGCAACATGCGACCACTGTG	CAAAATTTCCGAGATCCCGATTAAGCGCTTGGTGC	4922
Db	10063	CCTACGAACATGCGACCACTGTT	CCAAATGTGCCACAGATACCGTATTAAGGCACTTGTTG	10122
OY	4923	AACGCGCAGTTACGCGCACTT	TAAGTGAATCAAGGTCTGTCTCATGTGGAATTAAAC	4982
Db	10123	AAAAGGGCAGGTAATGCCCC	CGCTCAATTTGAGATCACTGTCAATGCTCTCGAGAGTTTGC	10182
OY	4983	CTTCAACATTAACAAAGATAGTGA	CGTCAAAATTTCCACACATCATCTTCTTACCAACAAG	5042
Db	10183	CTTCCACCAACCAAGATGATCA	TTTACCTGCAAAATTCACACATGTGTGTCCCTCCCAAAA	10242
OY	5043	TTAAATGTGCGGGTCCCTG	ATGTGCAAGGCATCTCMAAGCGGATTAACACATGCGCG	5102
Db	10243	TCAATATGTGCGCTCTCTT	GGAAATGATGCGCGCGCTCATGTGCAACTATACCTGTGAAG	10302

QY	5103	TTTTTGGCGGTGATACCCCTTTCATGTGTGGGAGGGGAGCGACAATGCTCTGTGAGACGTAGA	5162
Db	10303	TCCTTGGAGGGGATCTACCCCTTTATGTGTGGGAGGAGCGCAATGTTTTGGACAGTGA	1036
QY	5163	ACACACAACCTAGTGAAGGCGTACGCTCCAGATTCGCTCAGACTGCATTAATAGTACCAG	5222
Db	10363	ACACGCGATAGTGAAGGCGTACGCTCCAGATTTGTACAGAGATTTGGCGCTGTGACACGGCG	1042
QY	5223	TCGCATTAAGTTTCAACAGCTGCTCTGAAGTGGGCTGTGCTGTATGTATACGCAACA	5282
Db	10423	AGGCGATTAAAGTGCACACTGCCCGATGAAGTAGAGACTGCGTATTTGTGTACGGGAACA	1048
QY	5283	CCACCGGCGCACTTGATATCGTTTTTCATATGGCCGTACGGCAGGTTCTCAACGGGACTGA	5342
Db	10483	CTACCGATTTTCTTGAATGTGTACGTGAACGAGTCAACAGGAACGCTTAAACACTTGA	1054
QY	5343	AGTCATATGACGGGCGCATATCACGCGCTTTTACCCCTTGAACATTAAGTGTGTATCA	5402
Db	10543	TAATCATATGCTGACACCAATTTACAGATCTGTTTACGCCATTCGATCATTAAGTGTATCC	1060
QY	5403	GAAGGGGCTTGTTCACACTACGACTTCCCTGAGTATGAGCTATGAACCAAGACGCT	5462
Db	10603	ATCCGGGCTGTGTGACACTATGATCATTTCCCGAATATGAGCGATGAACCAAGACGCT	1066
QY	5463	TCGGCGATTTCAAGCATCTCTCGCTGTATGTACAGACATATGATAGCCCGCATGACATAC	5522
Db	10663	TTGGGACATTTCAAGCTACCTCTTGAATACCAAGGATCTCATGCGCAGCACACACATTA	1072
QY	5523	GGCTGTGAACCTCTTGTGTCAAGAACATCCAGTCCCTACACCCAGACAGTATCAGAGGT	5582
Db	10723	GGCTACTAAGCTTTCGCGCAAGAACGTGATGCTCCGTACACGACAGGCTCATACAGAT	1078
QY	5583	ATGAATGTGAGAGAACAACTCAGAGCAGCCCTGCAAGAAACAGACCAATTTGGATGTA	5642
Db	10783	TTGAGATGTGAAGAAACAACTCAGCGCGCCACATGCAAGAAACCGCACTTTGGGTGTA	1084
QY	5643	AAATTTGAATGAGAGCTCTGCGAGCGTCTTACTGTCTTACGGGCAATCCCTATCTCGA	5702
Db	10843	AGATTTGACGTAAATCCGCTCCGACCGGTGAGCTGTCTATACGGGAACATTTCCATTTCTA	1090
QY	5703	TTGACATCCCGATGACAGCTTTTGTGAGATCATGAGATCAACAAATTTTAAAGTTA	5762
Db	10903	TTGACATCCCGAAGCTGTCTTTATCAGGACATGAGATGACCACTGTGTCTACACATGCA	1096
QY	5763	GCTGCACAGTACAGACACTGCAATTTATTTCTGAGACTTTTGGTGTCTCTTAACATTACAGT	5822
Db	10963	AATGTGAAGTCAGTGAGTGCACATTATTCAGAGACTTGGGGGATAGGCCACCTGCACT	1102
QY	5823	ACAAAGCTGACAGGAGGAGCAATTTGTCAGTTCACTCCCACTCCACAGACAGCTGTTTGA	5882
Db	11023	ATGTATCCGAGCCGGAAGGTCAATGCCCGGTACATTTGCAATTCAGACAGCAACTCTCC	1108
QY	5883	AGGAAGCAGCAACACATGTGACTGCGGTAGGACACATTAACATTAATTTAAGCACAATGA	5942
Db	11083	AAGAGTCAACAGTACATGTCTCTGAGAAAGGAGCGGTGACAGTACACTTTAAGCACCCGGA	1114
QY	5943	GCCCAAGCAAGAAATTTTATAGTTTTCGCTATGCGGCAAGAAAGTCACTGCAATGCTGAAT	6002
Db	11143	GTCACAGGCGCAACTTATGTATGCTATGCTGTGTGGAAAGAAAGACAAACATGCAATGCAAT	1120
QY	6003	GTAACCAACCGGCGGACCACTAAATTTGAGAACCACTAAAGTGCACCAAGATTTCCAG	6062
Db	11203	GTAACCAACCAAGTGCACATATGCTGAGCAACCCGCAACAAAATGACCAAGAAATTTCAAG	1126
QY	6063	CGGCAAGTTCCAAAACATCTTGGAACTGGCTGCTTGCACGTGTTGGGGGAGCATATCCC	6122
Db	11263	CCGCGATCTCAAAAACATCAATGAGGTGGCGTGTGCTCTTTTGGGGCGGCTCGTCC	1132
QY	6123	TCATTTGTTGAGACTTATAGTGTGTGGCTGCTGACCTCATGCTTAAACACACAGTATAT	6182
Db	11323	TATTTATTTATGACATTATGATTTTTTGTCTTGACAGATGATGTCTACATGACACGAAGAT	1138
QY	6183	GACTGAGCGCGACACTGA-CATAGCGGTAAACATCGATGATCTTCGAGGAGCGTGT	6241

Db	11383	GACCGCTACGCCCCAATGATCCGACGACAAACTGATGTACTTCGAGGAACTGATGT	11442
Qy	6242	GCATATGCCACGCGCGCT	6261
Db	11443	GCATATGCAATCAGGCTGGT	11462

Search completed: November 16, 2003, 09:06:06
Job time : 380 secs

